DOCTOR OF PHILOSOPHY

The regulation and function of mRNA cap methylation in pluripotency and differentiation

Grasso, Laura

Award date:
2015

The regulation and function of mRNA cap methylation in pluripotency and differentiation

Laura Grasso

Supervisor: Dr. Victoria Cowling

This thesis is submitted for the degree of Doctor of Philosophy

College of Life Science

University of Dundee

14th October 2015

Declarations

I hereby declare that the following thesis is based on the results of investigation conducted by myself, and that this thesis is of my own composition. Work other than my own is clearly indicated in the text by reference to the researchers or their publications. This dissertation has not in whole, or in part, been previously presented for a higher degree.

.....

Laura Grasso

I certify that Laura Grasso has spent the equivalent of a least nine terms in research work in the College of Life Science, University of Dundee, and that she has fulfilled the conditions of the relevant Ordinance and Regulations of the University of Dundee and is qualified to submit the accompanying thesis in application for the degree of Doctor of Philosophy.

.....

Dr. Victoria Cowling

Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Victoria Cowling for her help and guidance over the last years. Thanks Vicky for letting me explore this project and having offered me advices and support when I needed the most. Thanks because you made possible anything I wanted to do and I will always be grateful for this. A special thank goes to Dr. Marios Stavridis for introducing me into the stem cells word. I am also grateful to my funding the Wellcome Trust and to the Korner travel fellowships for their generous financial supports.

I am particularly thankful to the past and current members of the VC lab. I have learned something from each of you, both personally and scientifically. You are the best lab ever. A special thank goes to Olga Suska (in our lab), I am glad you will take care of “my babies” and I really enjoyed working with you. And I have to especially thank Dhaval Varshney for his friendship as wells as help and support over my PhD.

In Dundee I got not only a PhD but also and most importantly a family. I refer to the Dundonian family, Giuli, Claudia and Kasiu. You girls made every day matters and be unforgettable.

Last but not least special thanks go to my beloved family, thanks for always encouraging me to pursue and carry on also in the darkest days. All my achievements will be meaningless if not shared with you. Not enough words can describe the gratitude towards Fabio, this PhD is mine as yours. At the beginning you encouraged me to start this experience and always pushed me to do better. I am the luckiest in the world to have you with me.

Abstract

The synthesis of the N7-methylguanosine cap at the 5' end of pre-mRNA occurs co-transcriptionally and is catalysed by a series of enzymes including the N7 RNA methyltransferase (RNMT), which along with its recently discovered activating subunit RAM, methylates the cap. RAM, which contains an RNA binding domain, is required to promote RNMT activity both *in vitro* and *in vivo*. Although the biochemical function of RAM has been characterized, its biological relevance remains elusive to date. The addition of the cap moiety is a crucial event in gene expression as it affects several processes within the mRNAs life cycle including mRNA processing, stability and translation. In stem cells, every step of mRNA metabolism is tightly regulated to maintain the undifferentiated state, allowing the expression of pluripotency genes and the concomitant repression of the lineage-specific ones. Here, I describe a critical role for the mRNA cap methylation in the maintenance of pluripotency. RNMT and RAM are highly expressed in mESCs compared to differentiated cells. The reprogramming of MEFs to iPS totally restores the elevated expression levels of RNMT and RAM suggesting that high levels of the two proteins are a feature of pluripotent cells. Even more exciting, the same expression is conserved amongst species as also hESCs and hiPSCs exhibit high levels of RNMT and RAM compared to fibroblasts. So far, RNMT and RAM were described as a complex in all cells lines examined and it was assumed that are similarly regulated, instead surprisingly, during *in vitro* neural differentiation a specific reduction in RAM protein levels is observed. Gain- and loss-of-function studies have been employed to demonstrate that specifically high RAM levels are required for the maintenance of pluripotency. In fact, RAM depletion causes a major reduction in the methyl cap levels of important pluripotency factors,

ultimately resulting in a decreased of the protein levels. Therefore, RAM is found to function as modulator of RNMT activity, whereby it promotes cap methylation of fundamental transcripts required for the maintenance of ESCs pluripotency. I have also found that during differentiation RAM is down regulated post-transcriptionally, and therefore current studies are focused on investigating the role of RAM phosphorylation at Serine-36, which correlates with proteosomal degradation. Together the data corroborate previous findings about the methyl cap formation being a critical and regulated process within gene expression and propose a novel implication of this modification in the maintenance of pluripotency.

Table of Contents

List of Tables

Table of Figures

List of abbreviations

µ:	micron
4E-BP:	eIF4E-binding protein
AP:	alkaline phosphatase
APA:	alternative polyadenylation
AS:	alternative splicing
BMP:	Bone morphogenetic protein
Bp:	Base pair
BSA;	Bovine serum albumin
CBC:	Cap-binding complex
CDK:	Cyclin-dependent kinase
cDNA:	Complementary DNA
CE:	Capping enzyme
ChIP:	Chromatin Immunoprecipitation
CNS:	Central nervous system
CPSF:	Cleave by polyadenylation specific factor
CTD:	Carboxyl-terminal domain
Da:	Daltons
DMSO:	Dimethyl sulphoxide
DNA:	Deoxyribonucleic acid
Dox:	Doxycycline
DRB:	5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
DSIF:	DRB-sensitivity inducing factor
DSTT:	Division of signal transduction therapy
E1:	Ubiquitin-activating enzyme
E2:	Ubiquitin-conjugating enzyme
E3:	Ubiquitin-ligating enzyme
EB:	Embryoid bodies
EC:	Embryonal carcinoma
EC:	Embryonal carcinoma
eIF:	Eukaryotic initiation factor
Erk:	Extracellular-signal-regulated kinases
ESC:	Embryonic stem cell
Fg:	Flag-tag
Fgf4:	Fibroblast growth factor 4
Fox:	Forkhead box
G:	Gram
GFP:	Green fluorescent protein
GMP:	Guanosine monophosphate
GpppG;	Guanosine cap
GSK3:	Glycogen synthase kinase
GTP:	Guanosine triphosphatase
H3K27me3:	Histone H3 Lysine 27 trimethylation
H3K4me3:	Histone H3 Lysine 4 trimethylation
HA:	Hemagglutinin-tag
HEK293:	Human embryonic kidney 293 cells
hESC:	Human embryonic stem cell
HITS-CLIP:	High throughput sequencing following cross-linking immunoprecipitation
HMG:	High mobility group
Hr:	Hour
ICM:	Inner cell mass
ICM:	Inner cell mass
Id:	Inhibitor of differentiation
IP:	Immunoprecipitation

iPSC:	Induced pluripotent stem cell
IRES.	Internal ribosome entry site
KLF4:	Kruppel-like factor 4
KO:	Knockout
L:	Litres
LCR:	Locus control region
LIF:	Leukemia inhibitor factor
LIFR:	LIF receptor
M:	Molar
m^6A:	N6-methyladenosine
m7G or m^7GpppG	N7-methylguanosine cap
MEFs:	Mouse embryonic fibroblast
mESC:	Mouse embryonic stem cell
methionyl tRNA:	Met-$tRNA_i^{Me}t$
Min:	Minutes
miRNA:	Micro RNA
mRNA:	Messenger RNA
MT:	Methyltransferase
MTL:	Multiple transcription factors
n:	Nano
NELF:	Negative elongation factor
NLS:	Nuclear localisation site
NMD:	Nonsense mediated decay
NPC:	Nuclear pore complex
N-terminal:	N-terminal- amino terminal
ORF:	Origin of replication
OSKM:	Oct4 Sox2 Klf4 Myc
PAP:	Poly (A) polymerase
PARN:	Poly(A)-specific ribonuclease
PIC:	Pre-initiation complex
PML:	Promyelocytic leukemia protein
POU:	Pit, Oct and Unc
PRC2:	Polycomb repressive complex
PTC:	Premature termination codon
P-TEFb:	Positive transcription elongation factor b
PTM:	Post translational modification
RT-qPCR:	Quantitative reverse transcriptase polymerase chain reaction
RAM:	RNMT activating Mini protein
RBP:	RNA binding proteins
REST:	Re1 silencing transcription factor
RNA POL II:	RNA Polymerase II
RNA:	Ribonucleic acid
RNAi:	RNA-Interference
RNGTT:	RNA guanylyltransferase and triphosphatase
RNMT:	RNA methyltransferase
Rpm:	Revolution per minute
SAH:	S-adenosylhomocysteine
SAHH:	S-adenosylhomocysteine hydrolase
SAM:	S-adenosylmethionine
SDS-PAGE:	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
snRNA:	Small nuclear RNA
Sox:	SRY-related HMG box
TAF:	TBP associated-transcription factor
TBP:	TATA binding protein
TBP:	TATA binding protein
TFIID:	Transcription factor II D
TFs:	Transcription factors

THREX:	Transcription-Export
TLC:	Thin layer chromatography
TMG:	Trimethylguanosine cap
tRNA:	Transfer RNA
TSAA:	Translation state array analysis
TSS:	Transcription start site
U snRNA:	U small nuclear RNA
U snRNP:	Small nuclear ribonucleoprotein
Ub:	Ubiquitin
Utf1:	Undifferentiated embryonic cell transcription factor 1
UTR:	Untranslated region
UV:	Ultraviolet
v/v;	volume to volume

1 Introduction

1.1 The gene expression pathway

The molecular biology discipline is based on the central dogma formulated by Francis Crick in 1970 (Crick, 1970). He firstly introduced the concept of gene expression describing how the hereditary DNA information is transcribed into RNA molecule to be then translated into proteins. Although the flow direction of gene expression is well conserved between prokaryotes and eukaryotes, the regulation mechanism of this pathway varies among species and between different cell types of the same organism. It is extremely fascinating how cells in our body, despite possessing the same genetic information, locally and temporally express different genes that function as hallmarks of their identity and allow cells to fulfil different roles. This high degree of regulation is critical to constantly adjust the RNA and protein content according to the environmental cues and intracellular signalling, and alterations of this plasticity were shown to be associated with abnormal cell proliferation and cancer (Delgado & Leon, 2006). Although in eukaryotes the gene expression pathway is characterized by only three steps, regulation occurs at multiple levels resulting in a sophisticated process that finely tunes the expression of a specific gene product within cells.

Gene regulation may occur at: epigenetic, transcription, post-transcriptional and post-translational level (Orphanides & Reinberg, 2002). As a perfect machine, all the above events co-ordinately ensure the expression of certain genes and the repression of others. In Prokaryotes and archaea, the DNA transcription in RNA is accomplished by a single RNA polymerase but this task in eukaryotes is divided within three highly related enzymes: RNA polymerase I and III devoted to express ribosomal and non-coding RNAs respectively, whereas the RNA polymerase II transcribed all the genes encoding messenger RNAs (mRNAs)

(Roeder & Rutter, 1969). An RNA polymerase of the mitochondrial origin, named snRNAP-IV, has also been found to transcribe mRNA of certain mammalian protein coding genes in the nucleus (Kravchenko *et al.*, 2005).

1.1.1 Transcription initiation

Transcription of a specific DNA segment into mRNA represents the first step of gene expression. In eukaryotes, a multi subunit DNA dependent RNA polymerase (RNA Pol II) is responsible for the transcription of genes encoding mRNA, additionally to U small nuclear RNA (U snRNA) and micro RNA (miRNA). Transcription by RNA Pol II is preceded by the sequential and coordinated assembling of different transcription factors (TFs), resulting in a multi-protein complex generally referred as pre-initiation complex (PIC). Firstly, the TFIID that contains the TATA binding protein (TBP) binds the core promoter element, which dictates initiation and orientation of transcription (Hahn, 2004). Subsequently TFIIA and TFIIB are recruited to stabilise the core-promoter bound TFIID and lastly, the PIC is completed by the binding of TFIIE and TFIIH (Nikolov & Burley, 1997). In addition to these basal transcription factors, the transcription initiation complex may include further regulators such as the mediator complex and/or gene-specific transcription factors which activate or repress transcription in response to metabolic or environmental signals (Sikorski & Buratowski, 2009).

The largest subunit of RNA Pol II (Rbp1) is characterised by the presence of a specific structure on the Carboxyl-terminal domain (CTD) that consists in tandem heptad repeats with a consensus sequence of $Tyrosine_1$-$Serine_2$-$Proline_3$-$Threonine_4$-$Serine_5$-$Proline_6$-$Serine_7$. Although the CTD is an exclusive

and ubiquitous feature of RNA Pol II, the repeats length varies among species and considering that yeast presents 26 repeat and human 52, it is believed the length increases with the complexity of the genome (Liu *et al.*, 2010). The CTD extends from the core enzyme close to RNA exit channel and from this strategic position, the CTD functions as landing scaffold that actively and dynamically recruits different components of the RNA processing machinery through the different stages of transcription, including chromatin remodelling and capping enzymes (Bentley, 2005; Egloff & Murphy, 2008). This high degree of CTD flexibility is reached through changes in the phosphorylation status of the Serine-5 and Serine-2 of the tandem repeats.

RNA pol II is recruited to promoters in a non-phosphorylated version (Feaver *et al.*, 1991) however, following the assembly of the PIC, the CTD is subsequently phosphorylated by TFIIH-associated kinase CDK7 (cyclin-dependent kinase 7) at the Serine-5 (Lu *et al.*, 1991). This event is associated with promoter clearance and release of the mediator complex from RNA Pol II (Sogaard & Svejstrup, 2007). Chromatin immunoprecipitation (ChIP) followed by sequencing revealed that shortly after promoter clearance, RNA Pol II pauses at the promoter proximal site (Guenther *et al.*, 2007; Rahl *et al.*, 2010; Kim *et al.*, 2005). More precisely, when RNA Pol II reaches 20-60 bp after the transcription start site (TSS), two factors bind to it: 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitive induced elongation factor (DSIF) and negative elongation factor (NELF) causing the RNA Pol II to pause. DSIF consists of the two subunits Spt4 and Spt5 whereas NELF is composed of five polypeptides named NELF A-E (Wada *et al.*, 1998; Yamaguchi *et al.*, 1999). This pausing occurs also for genes thought to be transcriptionally inactive, suggesting that transcription is regulated at the elongation step rather than at

initiation. The negative effect of these factors is reversed by CDK9 or P-TEFb (the positive transcription elongation factor b), which by phosphorylating the CTD Serine-2 as well as the two factors, promotes DSIF and NELF release and thus favours transcription elongation (Peterlin & Price, 2006; Rahl et al., 2010). The scenario that emerges is that high level of phosphorylated Serine-5 are found at the TSS whereas high level of phosphorylated Serine-2 are found downstream the TSS and when RNA Pol II is actively engaged in elongation. It has been recently reported that RNA Pol II landscape is more complex than it was originally thought because other CTD residues (threonine 4 and tyrosine 1) can be phosphorylated or modified with other post translational modifications (Egloff & Murphy, 2008; Heidemann *et al.*, 2013). These features further amplifies the possible regulation mechanisms of RNA Pol II in the transcription process.

1.1.2 Pre-mRNA processing

Once the pre-mRNA is transcribed, three main events, catalysed by different protein complexes, are required for the pre-mRNA maturation: the addition of N7-methylguanosine cap, splicing and polyadenylation. The CTD complementary phosphorylation pattern of Serine-5 and Serine-2, temporally and spatially coordinates the recruitment of the different protein complexes to ensure fidelity and efficiency of gene expression (Hsin & Manley, 2012; Moore & Proudfoot, 2009). The first process consists in the addition of methylguanosine cap at the 5' end of an emergent pre-mRNA, and the mechanism and function of this modification will be later discussed.

Splicing is the second mRNA modification that occurs when the methyl-capped mRNA encounters the spliceosome complex, recruited by phosphorylated CTD (Braunschweig *et al.*, 2013). Splicing is a two-step process that provides first the removal of non-coding intronic sequences, or introns, and secondly the joining of exons fragments to generate a functional message. In human, five main small nuclear ribonucleoprotein particles (U snRNPs) U1, U2, U4, U5 and U6 constitute the spliceosome machinery along with a large number of other proteins (Hoskins & Moore, 2012). Evolutionary processes result in alternative splicing (AS) documented as differential use of splice site that generates two different mRNAs and thus two different proteins. The evidence that 95% of human multi-exon genes is alternatively spliced highlights how the production of different protein isoforms from a single gene represents a real strategy to expand the genomic coding capacity (Braunschweig et al., 2013; Fu & Ares, 2014).

Temporally, the last modification of pre-mRNA is the polyadenylation, which is a two-step reaction that occurs in all protein encoding mRNAs with the exception of histone transcripts. The pre-mRNA is first cleaved at a certain signal sequence and then the poly (A) polymerase (PAP) catalyses the addition of 200-250 adenosine residues upstream to the cleavage site (Di Giammartino *et al.*, 2011; Bentley, 2005). Also for the 3' process has been shown that phosphorylated Serine-2 of CTD is required to mediate the recruitment of the processing factors (Hocine *et al.*, 2010; Hsin & Manley, 2012). The polyadenylation step is required for mRNA stability, export, localization and translation efficiency. Alternative polyadenylation (APA) events have been reported and similarly to alternative splicing, they further increase the number of mechanisms to diversify gene expression control (Di Giammartino et al., 2011).

Following the maturation process, mRNA is transported into the cytoplasm to be translated. In eukaryotes, the physical and functional compartmentalization between nucleus and cytoplasm triggered the evolution of a nucleocytoplasmic machinery to tightly regulate the trafficking of RNAs and proteins between the two compartments. Different nuclear pore complexes (NPCs) are present on the nuclear membrane, and each has nuclear receptors that will specifically recognise nucleotide motifs in RNA cargoes allowing their transition. This is the case for miRNAs and tRNAs (Kohler & Hurt, 2007), whereas the mRNA due to the different length and structure, needs some adaptors to interact with its export receptors. Several adaptors have been identified among which TAP-15 complex, exon-junction complex, and also the cap binding complex that will be further discussed. Another complex involved in the mRNA nuclear export is the THO/TREX (TRanscription-Export) so called because the components of this complex are involved in both transcription elongation, transcript-dependent recombination and nuclear export (Hocine et al., 2010; Kohler & Hurt, 2007). In fact, it has been described a model where the TREX complex is recruited to the transcribing genes and travels with the polymerase during the transcriptional elongation step (Strasser *et al.*, 2002). These findings highlighted how the different steps involved in mRNA biogenesis are physically and functionally coupled, and failure of one process compromises the progression into the other step (Katahira, 2012).

1.1.3 Translation initiation

Following transcription and mRNA processing, mRNA functions as template for protein synthesis. Efficient translation of most mRNAs is dependent on the N7-methylguanosine cap moiety as the m7G is bound by the eIF4F complex which

consist in eIF4E, eIF4G and eIF4A. eIF4E specifically recognises the N7-methylguanosine cap and interacts with the scaffolding protein eIF4G, which presents binding sites also for the RNA helicase eIF4A, which unwinds otherwise inhibitory RNA secondary structures (Koromilas *et al.*, 1992) and for eIF3. The latter promotes not only the assembly of the 43S pre-initiation complex (PIC), which consists in 40S ribosomal subunit, the initiator methionyl tRNA (Met-tRNA$_i^{Met}$) and several translation initiation factors but also its recruitment to the mRNA. Once the PIC is engaged to mRNA, it scans for the AUG start codon that would base pair with Met-tRNA$_i^{Met}$. Subsequently, certain translation factors are removed and the 60S ribosomal subunit joins the 80S initiation complex proceeding into the elongation stage (Jackson *et al.*, 2010; Sonenberg & Hinnebusch, 2009). eIF4E function can be repressed by eIF4-BPs, which in an unphosphorylated state binds to eIF4E preventing its interaction with eIF4G (Sonenberg & Hinnebusch, 2009; Topisirovic *et al.*, 2011).

Additionally to the canonical scanning mechanism, an alternative cap independent mechanism of translation initiation has been described. A subset of mRNAs can engage with the 40S subunit via specific sequence within their 5' untranslated regions (UTRs) known as internal ribosome entry site (IRES). IRES were originally observed in virus mRNA but it is now established that eukaryotic mRNA contains IRES and they recruit the 40S subunit in proximity of the start codon, independently from the methyl cap structure. It has been suggested that the cap dependent translation may occur in optimal growth condition, whereas as soon cells face any kind of stress the IRES dependent translation replaces the canonical process (Komar & Hatzoglou, 2011; Spriggs

et al., 2010). Different IRES structures exist within cells and their function and mechanism remain unclear.

1.2 The N7-methylguanosine cap structure

It was 1975 when the N7-methylguanosine cap (m7GpppN or simply m7G) was identified at the 5' end of viral mRNA (Furuichi & Miura, 1975; Furuichi *et al.*, 1975a) and subsequently its presence was described also on HeLa mRNA (Furuichi *et al.*, 1975b). This structure, indicated as cap 0 (Figure 1.1a), is found in all eukaryotic RNA Pol II transcripts and on eukaryotic viral RNAs but evidence of its presence was not reported for bacterial or archael RNA (Shuman, 2002). The N7-methylguanosine cap consists of a guanosine cap, linked to the first transcribed nucleotide via the unusual 5'-> 5' triphosphate linkage (Shatkin, 1976), further methylated in N7 position. The characterization and function of the cap 0 will be further discussed.

Additionally to cap0, mRNA of high eukaryotes but not yeast (Furuichi & Shatkin, 2000), was found to be methylated at the 2'-O position of the first or second nucleotide identifying respectively cap 1 and cap2 structure (Figure 1.1b) (Furuichi et al., 1975b). The methyltransferase, hMT1, which is responsible for the methylation of cap 1 and hMT2, required for the cap 2 methylation, have been only recently characterised (Belanger *et al.*, 2010; Werner *et al.*, 2011). Evidence suggest that hMT1 may function within the nucleus whereas hMT2 is thought to be cytoplasmic. To date the dynamic of cap 1 and 2 formation and their biological function remain elusive. Moreover, small nuclear RNAs (snRNAs) exhibit an alternate methyl cap structure that carries methylations on the positions 2,2,7 of the terminal guanosine, generating a trimethylguanosine (TMG) cap (Reddy *et al.*, 1992).

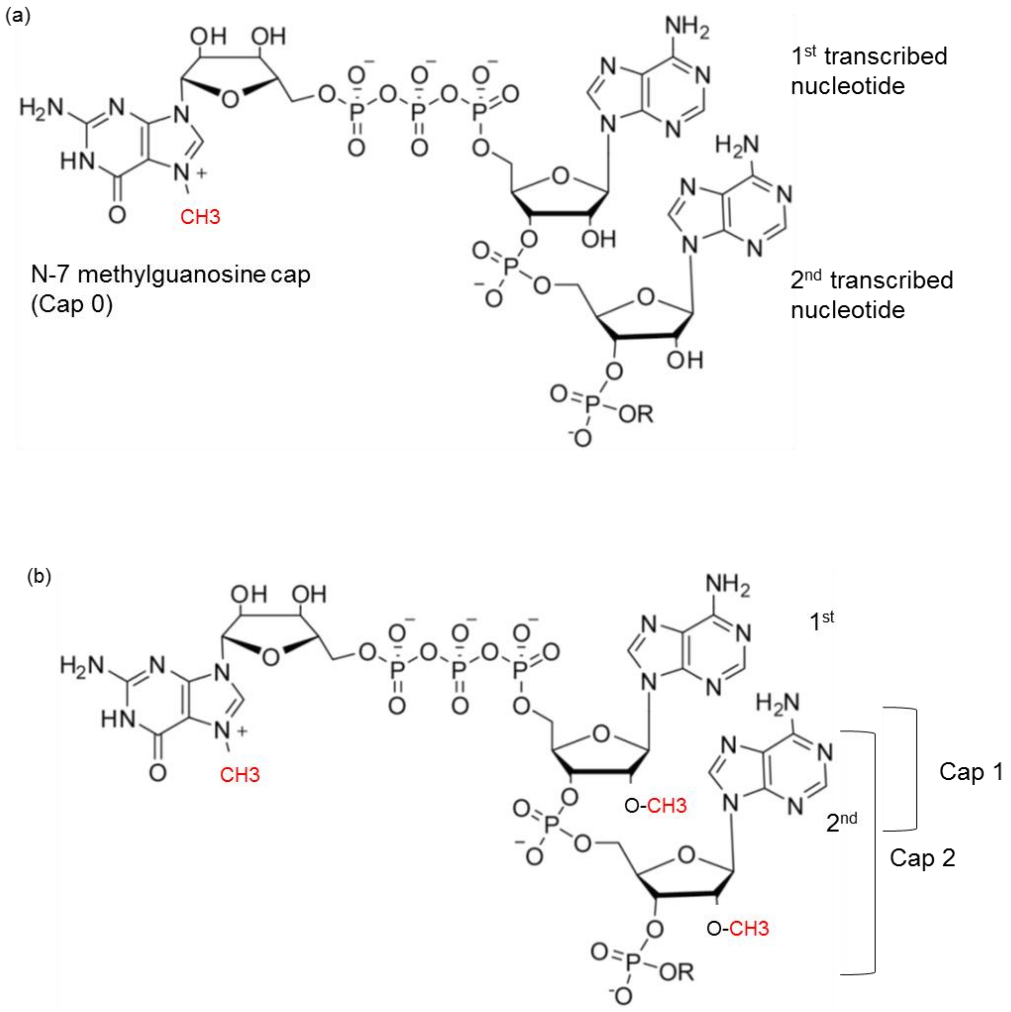


Figure 1.1: The structure of the N7-methylguoanosine cap (Cap0), Cap1 and Cap2.

(a) The N7-methylguaonsine cap (Cap 0) consists of an inverted guanosine group that is linked via a triphosphate bridge to the first transcribed nucleotide and is methylated at the N7 position. (b) In addition to the Cap O, the Cap1 structure is methylated at the 2'-O ribose of the first transcribed nucleotide and the Cap2 structure is further methylated at the 2'-O ribose of the second transcribed nucleotide.

1.2.1 Enzymology and mechanics of N7-methylguanosine cap

Three sequential enzymatic activities coordinate the formation of the N7-methylguanosine cap. The process starts when the RNA triphosphatase hydrolyses the γ-phosphate from the triphosphate end of the pre-mRNA generating a dephosphorylated 5' end (pppN → ppN). The resulting diphosphate is then modified by the addition of a guanosine monophosphate (GMP) known as guanosine cap (ppN → GpppN), catalysed by the guanylyltransferase. Lastly, the process terminates when the RNA methyltransferase transfers a methyl group from the methyl donor S-adenosylmethionine (SAM), to the guanosine cap in the N7 position (GpppN → m7GpppN) (Figure1.2) (Furuichi & Shatkin, 2000). Cap methylation is a well conserved mechanism that occurs in all eukaryotes (also in eukaryotes virus) and the majority of our understanding derives from extensive studies performed in yeast.

In lower eukaryotes such as yeast, the catalysis of the three reactions to form the methyl cap is assigned to three independent polypeptides (Mao *et al.*, 1995; Shibagaki *et al.*, 1992; Tsukamoto *et al.*, 1997). On the contrary, in metazoa and mammals evolution resulted in a unique bi-functional enzyme called RNGTT or capping enzyme (CE), which harbours the triphosphatase activity on its N-terminal domain (1-210) and the guanylyltransferase activity on its C-terminal domain (211-597). 25 amino acids residues are predicted to form a flexible loop that merges the two domains together (Pillutla *et al.*, 1998a; Yamada-Okabe *et al.*, 1998; Yue *et al.*, 1997; Chu *et al.*, 2011). However, the methyltransferase activity resides in a distinct protein known as RNA

methyltransferase (RNMT) (Pillutla et al., 1998a; Pillutla *et al.*, 1998b; Saha *et al.*, 1999; Tsukamoto et al., 1997).

The first step of the methyl cap reaction is catalysed by the RNA triphosphatase and the sequence analysis revealed that it is not conserved amongst eukaryotes (Changela *et al.*, 2001; Lima *et al.*, 1999). The human triphosphatase RNGTT has been described as belonging to the cysteine phosphatase superfamily and the specificity towards its substrate, the triphosphate group, is accomplished by the deep active site that can accommodate only the triphosphate group but it is inaccessible to diphosphate or monophosphate.

The guanylyltransferase is responsible for the second step of the reaction and catalyses the addition of the guanosine cap to the diphosphate mRNA. Contrary to the triphosphatase enzyme, the guanylyltransferase are structurally related in eukaryotes (Chu et al., 2011; Fabrega *et al.*, 2003; Gu *et al.*, 2010; Hakansson *et al.*, 1997; Hakansson & Wigley, 1998). The addition of the guanosine cap occurs in two steps: firstly the GTP binds a Lysine within the enzyme active site generating an intermediate enzyme-lysine-GMP conjugate. The GTP binding induces the guanylyltransferase to assume an open conformation able to release pyrophosphate and to ligate GMP to the diphosphate RNA (Fabrega et al., 2003; Hakansson et al., 1997; Hakansson & Wigley, 1998).

The last step in the methyl cap synthesis is the methylation of the guanosine cap mediated by the RNA methyltransferase enzyme. RNMT catalyses the transfer of a methyl group from the methyl donor S-adenosylmethionine (SAM)

to GpppRNA resulting in the m7GpppRNA and releasing the by-product of the reaction S-adenosylhomocysteine (SAH). Studies of the crystal structures revealed how the active site, which is conserved amongst eukaryotes, contains two distinct ligand-binding pockets to accommodate the methyl donor SAM and the methyl acceptor guanosine cap. Moreover, the crystal structure demonstrated as RNMT does not physically interact with either the methyl donor or the methyl acceptor, suggesting that RNMT role is to orientate the substrates in proximity to facilitate the methyl transfer (Fabrega *et al.*, 2004).

Apart from the triphosphatase, the enzymes required for the formation of the methyl cap are functionally conserved from yeast to human. The growth defect of yeast cap guanylyltransferase or cap methyltransferase deletion mutants can be totally rescued by their mammalian orthologues (Saha et al., 1999; Yamada-Okabe et al., 1998; Yue et al., 1997). In addition, it has been reported that the human and yeast capping enzymes, are essential for cell viability (Chu & Shatkin, 2008; Shafer *et al.*, 2005). There is no redundant triphosphatase, guanylyltransferase and methyltransferase which can compensate the loss of RNGTT and RNMT. All together these observations highlight the important role that these enzymes play in the gene expression pathway.

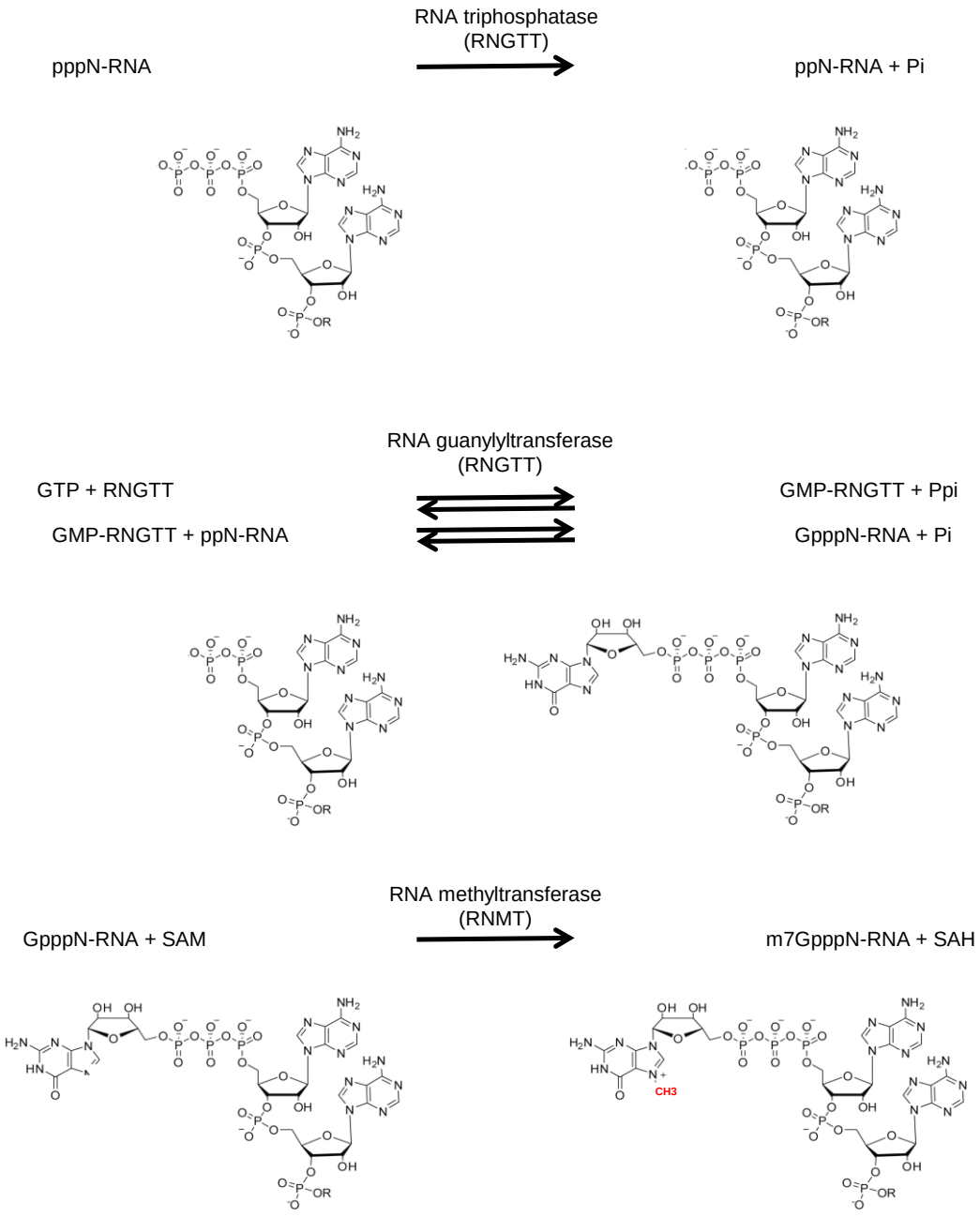


Figure 1.2: Reactions that are required for N7-methylguanosine cap synthesis. The synthesis of the N7-methylguanosine cap is described in the main text. The activities responsible for the reaction catalysis are indicated next to the arrows. The name of the actual enzyme in mammals is indicated in brackets. RNGTT contains both the tryphosphatase and the guanylyltransferase activity while RNMT harbours the methyltransferase activity. Contrary to the methyltransferase and the triphosphatase, the guyanylyltransferase reaction is reversible.

1.2.2 The methyltransferase complex: RNMT and RAM

Homo Sapiens RNA methyltransferase consists of 576 amino acids with a molecular weight of 64 kDa. RNMT is structurally divided in two domains: N-terminal domain (1-120) and C-terminal domain (120-476). The N-terminal domain was initially speculated to have a regulatory mechanism whereas the catalytic activity resides in the C-terminal domain (Saha et al., 1999). Immunofluorescence studies on different RNMT truncated constructs demonstrated that RNMT is a nuclear protein and three redundant nuclear localisation site (NLS) were found at residue 80, 103 and 126 (Aregger & Cowling, 2013; Shafer et al., 2005; Wen & Shatkin, 2000). Yeast two-hybrid experiments combined with *in vitro* pull down assay showed that the nuclear localization is mediated by the binding with Importin-α. Importin-α is the adaptor protein that associating with Importin-β, which interacts with the NPCs (Lange *et al.*, 2007), allows RNMT translocation into the nucleus (Shafer et al., 2005; Wen & Shatkin, 2000). Moreover it has been shown that Importin-α induces an increase in cap methylation of up to 10 fold possibly enhancing specific binding of RNMT to methyl capped RNAs whereas Importin-β was shown to prevent this stimulation (Wen & Shatkin, 2000). The N-terminal domain has recently been described mediating the recruitment of RNMT to the TSS (Aregger & Cowling, 2013). The catalytic C-terminal domain, is conserved in sequence, function and structure in all mRNA cap methyltransferases.

In 2011, the C-terminal domain of RNMT was found physically interacting with a small protein named RNMT Activating Miniprotein (RAM) (Gonatopoulos-Pournatzis *et al.*, 2011).

Homo sapiens RAM consists of 118 amino acids with a molecular mass of 14.3 kDa. RAM is only present in vertebrates and it is well-conserved in higher eukaryotes. Three functional domains were identified: the N-terminal activation domain (amino acids 1-55) is well conserved among vertebrates and it directly interacts with the catalytically active RNMT domain. The central RNA-binding domain (56-90) is enriched in positively charged Arginine and Asparagine residues (Figure 1.3a), which are characteristic of proteins with a RNA binding activity (Bayer *et al.*, 2005). The RAM C-terminal domain (amino acids 91-118) is not well conserved and is highly enriched in Glutamine, Tyrosine and Proline (QYP) amino acids. Recent findings in our lab established that this domain is required for the nuclear import of RAM as it contains two PY-NLSs. PY motifs are recognised by karyopherin β2 that binds to cargoes and target them to the NPCs. Therefore, similarly to other RNA-binding proteins, RAM via direct interaction with karyopherin β2 is imported into the nucleus where it interacts with RNMT (Gonatopoulos-Pournatzis & Cowling, 2014).

Studies *in vivo* with MG132 treatment showed that RNMT and RAM protect each other from proteasome degradation (Gonatopoulos-Pournatzis et al., 2011). Moreover, methyl cap immunoprecipitations demonstrated that RAM depletion caused an impairment in the level of capped endogenous transcripts. Combining these observations together RAM was described as an obligate component of the mammalian cap methyltransferase complex because required for promoting RNMT activity both *in vitro* and *in vivo*. Interestingly, the RNA-binding motif is not required for RNMT activation suggesting that RAM binding may induce a conformational change in RNMT that results in an increased methyltransferase activity.

To date, no crystal structures of RNMT and RAM complex are currently available, nonetheless studies in our laboratory suggest that RAM may facilitate the recruitment of the methyl donor SAM to RNMT binding pocket (Varshney et al., unpublished). However, the exact mechanism of how RAM functions or how is regulated have not been uncovered yet.

1.3 N7-methylguanosine cap synthesis occurs co-transcriptionally

As described earlier, the synthesis of N7-methylguanosine cap is the first pre-mRNA processing event. Several studies showed it occurs co-transcriptionally, possibly as soon as the transcript emerges from the exit channel of RNA Pol II (Figure 1.3b). In fact, different studies showed that transcripts between 20 and 80 nucleotides in length do present the methylguanosine cap structure (Coppola *et al.*, 1983; Mandal *et al.*, 2004; Rasmussen & Lis, 1993). First evidence to suggest that the cap methylation is coupled with the transcription process came from studies *in vitro* showing that formation of the methyl cap proceeds faster when RNA is associated in elongation complex compared to free RNA. Additionally, it was observed that the S-adenosylhomocysteine (SAH), the by-product of the methylation step, inhibits the *in vitro* transcription (Jove & Manley, 1984; Moteki & Price, 2002). Moreover, when CTD truncated version of RNA Pol II was used for the transcription, an impairment of capping process was observed suggesting that the concomitance in transcription and capping events is very likely permitted by the physical interaction between CTD of RNA Pol II and the capping machinery (McCracken *et al.*, 1997).

The model of co-transcriptional N7-methylguanosine cap formation is further corroborated by studies on association of the capping machinery with chromatin and RNA Pol II. In fact, ChIP analysis showed that particularly in human, RNGTT and RNMT are found at the TSS (Glover-Cutter *et al.*, 2008; Guiguen *et al.*, 2007; Komarnitsky *et al.*, 2000; Schroeder *et al.*, 2000) but also in the gene body and at the 3' end (Glover-Cutter et al., 2008).

In yeast, both guanylyltransferase and methyltransferase directly bind to the phosphorylated CTD, while the triphosphatase is recruited via binding of the guanylyltransferase (Cho *et al.*, 1997; McCracken et al., 1997). In mammals, a direct interaction of the RNGTT with phosphorylated CTD was also reported (Ho & Shuman, 1999; Ho *et al.*, 1998; Yue et al., 1997) but contrary to yeast, not direct association was observed between RNMT and RNA Pol II (Shatkin & Manley, 2000). However, recent findings in our laboratory showed that DRB treatment, the RNA pol II kinase inhibitor, reduced RNMT recruitment to promoters (Aregger & Cowling, 2013).

Subsequently, studies in human clarified that Serine-5 phosphorylation is not only required for the recruitment of RNGTT to TSS but also promotes the formation of the enzyme-GMP intermediate, favouring the capping reaction (Ho & Shuman, 1999; Wen & Shatkin, 1999). Recent biochemical data in yeast revealed that in addition to binding Serine-5 phosphorylated CTD, the capping enzyme requires an intact RNA 5'-triphosphate to efficiently bind to RNA pol II. The strategic position of capping enzyme allows an instant recognition of exiting RNA 5' end and the guanylation reaction, combined with the decrease in Serine-5 phosphorylated CTD, leads to complete release of capping enzyme (Martinez-Rucobo *et al.*, 2015).

Furthemore, it was observed that capping enzyme interacts also with DSIF subunit Spt5, and this association additionally to stimulate capping activity also relieves transcription repression by the negative elongation factor. This data are consistent with a role of capping enzyme in elongation checkpoint control during promoter clearance (Mandal et al., 2004).

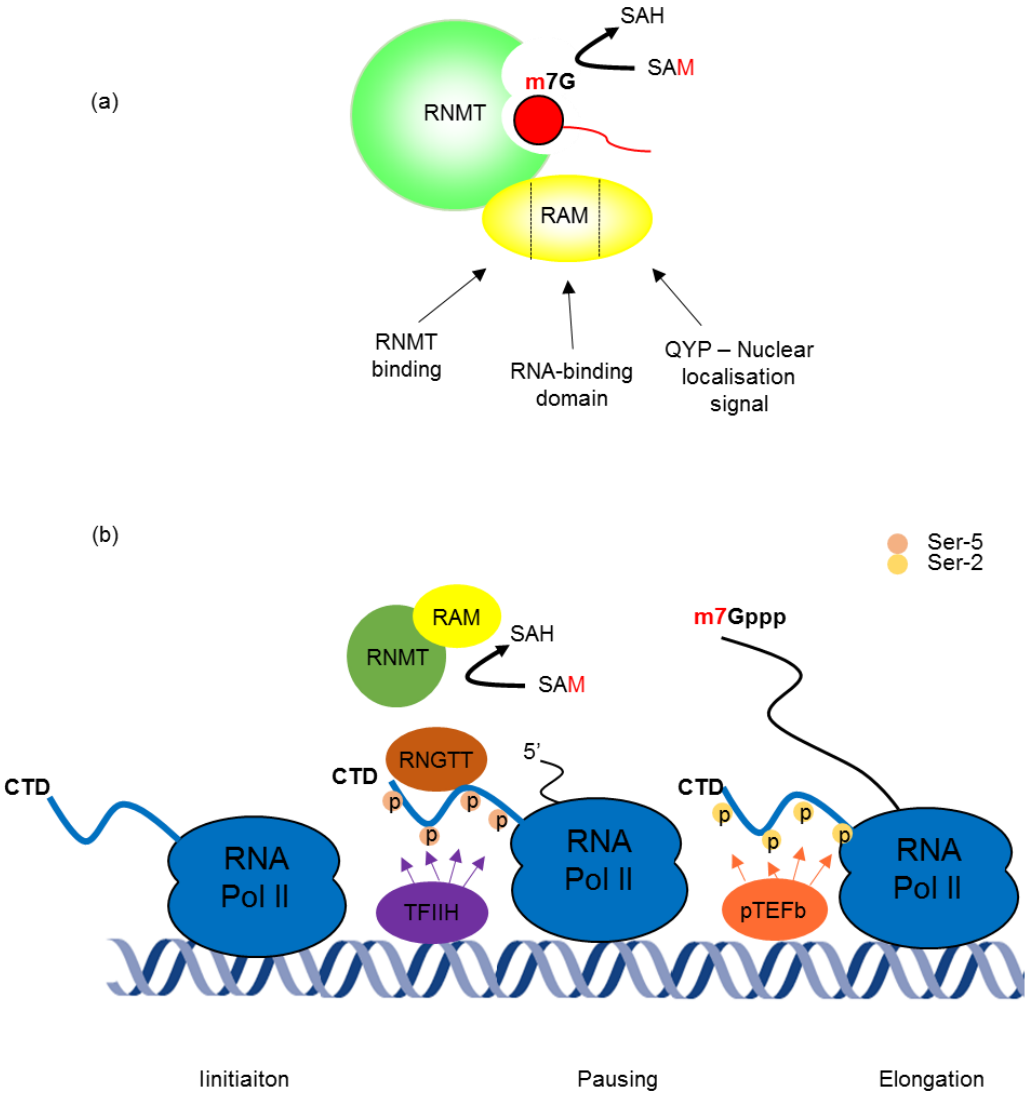


Figure 1.3: RNMT forms a complex with RAM and together with RNGTT provide the synthesis of the N7-methylguanosnie cap co-transcriptionally.
(a) As pre-mRNA emerges from RNA Pol II, RNGTT mediates the addition of the guanosine cap to the 5' end of mRNA. Then RNMT and RAM complex catalyses the methylation of the guanosine cap. S-adenosylmethionine (SAM) serves as methyl donor in the reaction. The C-terminal domain of RNA methyltransferase (RNMT) interacts with the N-terminal domain of RAM. RAM is a RNA-binding protein that recruits RNA to RNMT. (b) The CTD domain of RNA Pol II is strategically placed next to the RNA exit channel. The CTD, being sequentially phosphorylated on different residues, functions as scaffold and recruits RNA processing complexes. When RNA Pol II is recruited to promoter the CTD is hypophosphorylated, but shortly after initiation, TFIIH-associated kinase phosphorylates the CTD on Serine-5. As RNA Pol II elongates, Ser-2 is increasingly phosphorylated by CDK9 or P-TEFb, while Ser-5 phosphorylation is gradually removed by phosphatases. The recruitment of the capping machinery is dependent on Serine-5 phosphorylated RNA Pol II.

1.4 Regulation of the N-7 methylguanosine cap

Since the methyl guanosine cap was discovered, it was assumed to be a constitutive modification of all the RNA Pol II transcripts. On the contrary, increasing number of evidence over the last decades shed the light on possible regulation of this important step during mRNA maturation. Pioneering studies demonstrated that the transcription factors c-Myc and E2F1 promote the N7-methylguanosine cap of specific transcripts (Cole & Cowling, 2009; Cowling & Cole, 2007). The proto-oncogenes Myc and E2F1 are upregulated in response to growth factors and they are required for cell growth and proliferation. The Myc family comprises c-Myc, N-Myc and L-Myc (Meyer & Penn, 2008).

c-Myc has been recently described as an amplifier of the gene expression programme, resulting in increased levels of transcripts and thus in protein synthesis (Cowling, 2010; Lin *et al.*, 2012; Nie *et al.*, 2012). Initially it was observed that N-Myc increases translation of its target genes in a post-transcriptional mechanism without upregulating their mRNA levels (Cowling & Cole, 2007). Further experiments revealed that N-Myc increases the recruitment of TFIIH, which contains CDK7 that will phosphorylate CTD Serine-5 of RNA Pol II. As previously discussed, this phosphorylation event promotes capping machinery recruitment and activity (Cowling & Cole, 2007). In fact, experiments with an anti-methylguanosine cap antibody showed elevated N7-methylguanosine cap levels upon N-Myc expression. Additionally, it has been demonstrated that c-Myc promotes the methyl cap formation by relieving the by-product of the process. Following the methylation reaction, SAM, the methyl donor, is converted to S-adenosylhomocysteine (SAH) and elevated levels of

SAH inhibit RNMT activity. The loop is alleviated by the presence of S-adenosylhomocysteine hydrolases (SAHH) that removes the SAH allowing the reaction to carry on (Chiang *et al.*, 1996). It has been shown that c-Myc upregulates SAHH, which reduces SAH thus promoting RNMT activity (Figure 1.4a). c-Myc induced cap methylation is an essential mechanism by which c-Myc fulfils its proto-oncogene effect, promoting protein synthesis, cell proliferation and cell transformation (Fernandez-Sanchez *et al.*, 2009). Similarly to c-Myc, it has been reported that also E2F1 induces upregulation of the methylguanosine cap levels (Cole & Cowling, 2009) via promoting RNA Pol II CTD phosphorylation (Figure 1.4b) (Aregger & Cowling, 2013). Although c-Myc does not directly promote RNMT activity, it has been reported that RNMT overexpression induces cell transformation equivalently to c-Myc effect. The same study showed that methyl cap levels of oncogene Cyclin D1 increased in response to RNMT overexpression (Cowling, 2010). In light of these observations, the mRNA cap methylation emerges as a more sensitive process within gene expression than it was originally thought. Indeed the process can be regulated by environmental clues and internal signalling and the capping machinery need to be expressed within a certain threshold to avoid deleterious consequences. Thus the methyl cap formation may represent a critical step to directly and rapidly regulate the expression of specific genes.

Furthermore, latest evidence reported that in yeast, amino acids or glucose starvation impairs the mRNA cap methylation process resulting in unmethylated capped RNAs (Jiao *et al.*, 2010).

Recent studies showed that aberrantly capped RNAs may be generated also in normal growth condition and a novel class of endonucleases, which includes Rai1 and Dxo1, was first identified in yeast. Subsequently, the human orthologue DXO was found to specifically recognise and mediate the decay and clearance of incompletely capped mRNAs (Chang *et al.*, 2012; Jiao *et al.*, 2013; Jiao et al., 2010). These observations led to reformulate the assumption that mRNA cap methylation not necessarily always proceeds to completion. In fact, from these studies, it emerges that within cells aberrantly capped pre-mRNAs may be specifically degraded by particular enzymes that operate a surveillance mechanism to avoid the accumulation of potential deleterious defectively capped pre-mRNAs.

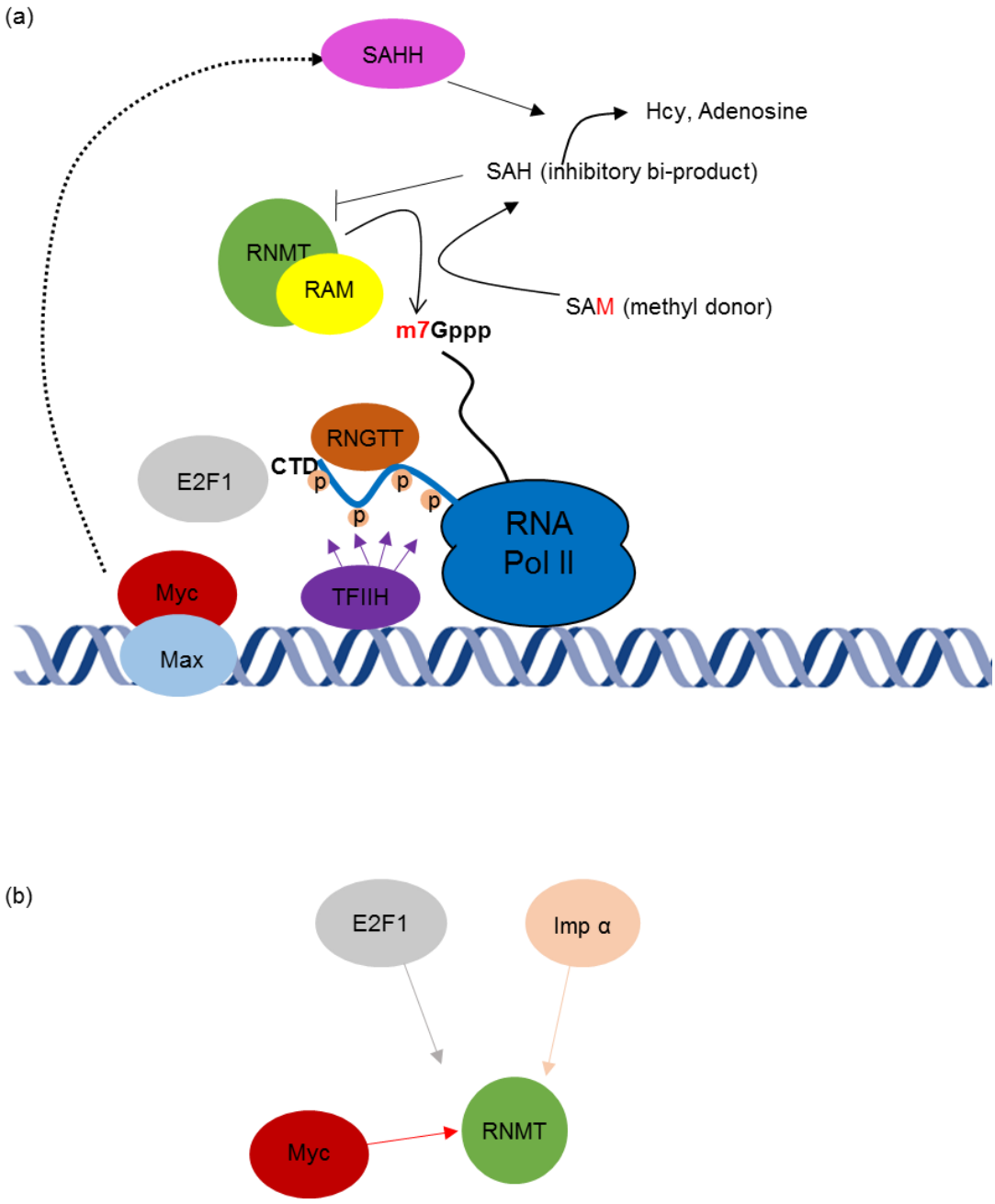


Figure 1.4: c-Myc and E2F1 upregulate N7-methylguanosine cap synthesis.
(a) Efficiency of cap methylation is increased by stimulation of RNMT activity by c-Myc and E2F1. Both factors promote methyl cap formation by increasing the recruitment of TFIIH, responsible for the phosphorylation of Serine-5 CTD. Moreover, c-Myc, which dimerizes with Max, promotes SAHH activity, which removes the inhibitory by-product of the methylation reaction. (b) Additionally, RNMT activity is regulated by Importin-α, which interacting with RNMT increases the cap methylation. This stimulation is inhibited by Importin-β.

1.5 N7-methylguanosine cap in gene expression

The synthesis of N7-methylguanosine cap moiety at the 5' end of transcripts is required for numerous processes within the gene expression pathway. Some examples include transcription, pre-mRNA processing, RNA stability, mRNA export and translation. The ability of m7G to accomplish so many different functions relies on its interaction with several binding proteins, amongst which, the cap binding complex (CBC) and eukaryotic translation initiation factors 4E (eIF4E) are the best characterised (Topisirovic et al., 2011). Additionally, other interacting partners have been reported such as the poly(a) binding protein C (PABPC), the RNA binding proteins Pumillo 2 (Pum2) and the exon junction component Y14 (Cao *et al.*, 2010; Chuang *et al.*, 2013; Khanna & Kiledjian, 2004), but their functional role remains unclear.

1.5.1 eIF4E and CBC

eIF4E was identified and purified as interacting partner of the m7G both in yeast and mammals (Altmann *et al.*, 1985; Altmann *et al.*, 1987; Sonenberg *et al.*, 1978; Sonenberg *et al.*, 1979). Its specific association with the methylated cap structures was lately confirmed by crystallography studies (Marcotrigiano *et al.*, 1997; von der Haar *et al.*, 2004). The interaction and binding between eIF4E and m7G clarified the importance of the methyl cap for a correct translation process. In fact, as previously discussed, the eIF4E, eIF4G and eIF4A are all part of the eIF4F complex, the role of which is to facilitate the engagement of the 43S pre-initiation complex with the mRNA (Sonenberg & Hinnebusch, 2009). The process starts when the scaffold polypeptide eIF4G interacts with

eIF4E, and this induces an allosteric change that increases the affinity of eIF4E to specifically bind the m7G (Gross *et al.*, 2003; Haghighat & Sonenberg, 1997). Therefore, the eIF4E is essential for cap-dependent translation since it firstly mediates the recruitment of the translation machinery onto the 5' end of mRNA. Considering these evidence, it does not surprise that overexpression of eIF4E was found to promote tumour formation in mice and in a plethora of human malignancies (Lazaris-Karatzas *et al.*, 1990; Ruggero *et al.*, 2004; Topisirovic et al., 2011).

CBC was found in HeLa extracts as heterodimer consisting of two subunits named according to their molecular weight Cbp20 and Cbp80. Competition experiments with vary cap analogues showed that the CBC binds specifically to the m7GpppRNA (Izaurralde *et al.*, 1995a; Izaurralde *et al.*, 1994). Moreover, RNA bind shift assay showed that both subunits synergistically bind RNA since none of them can bind the m7G as monomer (Izaurralde et al., 1995a; Izaurralde et al., 1994; Kataoka *et al.*, 1995). Crystal structures studies revealed that Cbp80 contains a nuclear localization site (Izaurralde *et al.*, 1995b) whereas the m7G binding pocket resides into the Cbp20 subunit. It was suggested that Cbp80 induces conformational changes required for Cbp20 to bind the m7G (Calero *et al.*, 2002; Mazza *et al.*, 2001; Mazza *et al.*, 2002). In mammals, depletion of CBC causes decrease in cell proliferation (Narita *et al.*, 2007; Pabis *et al.*, 2013). Although eIF4E and CBC fulfil different roles, they exhibit similar molecular structure of the cap binding pocket characterized by aromatic amino acids that specifically accommodate the m7G. Biophysical assay demonstrated CBC has 25% higher affinity for m7G than eIF4E (Worch *et al.*, 2005).

1.5.2 The N7-methylguanosine cap stabilises mRNA

One of the multiple functions of the m7G is the ability to stabilise the mRNA. Evidence to sustain this hypothesis were first collected when uncapped or capped RNA were microinjected into *Xenopus leavis* oocytes and the methyl cap was shown to be more stable than the uncapped version (Furuichi *et al.*, 1977; Green *et al.*, 1983). Furthermore, it was observed that ApppG was stable as the m7GpppG and that incubation of m7G analogues, which compete for binding to cap-binding proteins, did not destabilise the capped transcripts. All together these observations led to the conclusion that the methyl cap acts as a blocking structure sufficient to protect the mRNA from exoribonucleolytic degradation (Inoue *et al.*, 1989; Murthy *et al.*, 1991; Shimotohno *et al.*, 1977). RNA stability, a part from the cap itself, is strongly increased also by the consequent interaction with the cap binding proteins. In fact, several studies revealed that CBC and eIF4E, showing high affinity for the m7G, not only compete with decapping enzymes to bind to m7G (Grudzien *et al.*, 2006; Jiao et al., 2013; Schwartz & Parker, 2000) but also inhibits activity of poly (A)-specific ribonuclease (PARN), which catalyses the deadenylation step in mRNA decay (Balatsos *et al.*, 2006; Gao *et al.*, 2000). Noteworthy, and in contrast to the capping reaction, the methylation is not reversible (Shatkin, 1976) and considering that there are specific decapping enzymes degrading unmethylated capped RNAs, the presence of the methyl group further protects the cap structure preventing formation of aberrantly capped mRNAs and thus their degradation.

1.5.3 The role of N7-methylguanosine cap in transcription

ChIP studies found both Cbp20 and Cbp80 subunits at the 5' of genes as well as within the gene body (Glover-Cutter et al., 2008; Lahudkar *et al.*, 2011; Listerman *et al.*, 2006; Narita et al., 2007; Zenklusen *et al.*, 2002). Moreover, the evidence that only the heterodimer was bound to chromatin, strongly suggests that the CBC recruitment is mediated by m7G (Lahudkar et al., 2011; Wong *et al.*, 2007). In mammal cells, CBC which binds the pre-mRNA concomitantly to RNA Pol II pausing, was found to recruits P-TEFb which promotes transcription elongation as previously described. CBC depletion was indeed found to reduce levels of phosphorylated Serine-2 CTD, accompanied by accumulation of RNA Pol II in gene bodies (Lenasi *et al.*, 2011).

Additionally to CBC-mediated effects, also RNGTT was found to promote transcription elongation *in vitro*. It was showed that as soon as RNA Pol II pauses at the promoter, RNGTT is recruited to the transcription complex not only by phosphorylated Serine-5 CTD but also by DSIF subunit Spt5. The interaction between RNGTT and Spt5 rescues the NELF repressive effect restoring transcriptional elongation (Mandal et al., 2004; Pei & Shuman, 2002). These observations directly couple pre-mRNA capping with transcription elongation.

1.5.4 The role of N7-methylguanosine cap in splicing

The m7G plays also a pivotal role in pre-mRNA splicing as shown by incubation of *in vitro* uncapped or methyl capped transcripts with HeLa cell extracts.

These experiments demonstrated that the m7G moiety was required for efficient splicing event (Edery & Sonenberg, 1985; Izaurralde et al., 1994; Konarska *et al.*, 1984; Ohno *et al.*, 1987; Patzelt *et al.*, 1987). In fact, although both capped and methyl capped RNA were efficiently spliced, the addition of SAH, which inhibits the methylation reaction, inhibited only the splicing of capped RNA (Edery & Sonenberg, 1985; Ohno et al., 1987). Moreover, it was shown that incubation with the cap analogues m7GpppG and m7Gppp inhibited splicing more efficiently than the GpppG. These evidence combined together strengthened the hypothesis that the effect of cap methylation on splicing is mediated by a methylguanosine cap binding protein. In fact, immunodepletion of CBC showed that it is required for efficient pre-mRNA splicing and recent studies in mammals revealed that CBC is required for the recruitment of the spliceosome machinery (Pabis et al., 2013). Despite initial studies found that only the splicing of the 5' proximal intron was m7G dependent, latest findings reported that m7G and CBC promotes removal of downstream introns as well (Jiao et al., 2013). Recent evidence described that CBC, is also involved in the alternative splicing as it is able to recruit the alternative splicing factors SRSF1 (Lenasi et al., 2011).

1.5.5 The N7-methylguanosine cap and 3' processing

The hypothesis that m7G influences also the addition of poly (A) tail was based on experiments where *in vitro* transcribed mRNAs were injected into *X. leavis* oocytes, and the presence of the m7G moiety strongly enhanced the cleavage at the poly (A) site (Georgiev *et al.*, 1984).

Further experiments in mammals confirmed that methylguanosine capped transcripts are more efficiently cleaved than guanosine capped or uncapped transcripts (McCracken et al., 1997). Moreover, the addition of the cap analogue inhibits 3' end cleavage suggesting that also in this step, a cap binding protein could mediate the m7G effect on the RNA 3' processing. CBC was indeed found being involved in the process because following CBC depletion, HeLa extracts exhibited a reduced mRNA 3' end cleavage, and the effect was rescued by addition of recombinant CBC (Flaherty *et al.*, 1997). This confirmed that m7G moiety and CBC are required for correct pre-mRNA 3' processing although the exact mechanism remains unclear (Cheng *et al.*, 2006; Kohler & Hurt, 2007).

1.5.6 The N7-methylguanosine cap and mRNA export

After the maturation process, mature mRNA is translocated into the cytoplasm to be translated and this process was also revealed to be highly dependent on the m7G. In fact, it was shown that when methylguanosine cap and uncapped mRNAs were injected into *X. laevis*, the export of m7G mRNAs was more efficient compared to the uncapped RNAs. The potential role of m7G in mRNA export was validated by immunoprecipitation experiments were the Cbp80 subunit of CBC was found interacting with the Aly component of the TREX nuclear export complex (Cheng et al., 2006; Kohler & Hurt, 2007). It was found that the efficient recruitment of TREX to mRNAs is both cap and splicing dependent (Cheng et al., 2006). According to the RNA length other CBC-dependent mRNA export pathways have been identified (McCloskey *et al.*, 2012).

Furthermore, an eIF4E-mediated export of gene specific mRNAs was described, as experiments showed that an eIF4E mutant incapable to bind m7G was unable to export mRNA. Consistently to this, the addition of promyelocytic leukemia protein (PML), which reduces eIF4E affinity for the cap, also inhibited eIF4E-mediated mRNA export (Cohen *et al.*, 2001; Culjkovic *et al.*, 2006). Therefore, the eIF4E mRNA export seems to be methyl cap–dependent and is independent of its function in translation (Culjkovic et al., 2006). It is proposed that eIF4E mRNA export could provide a rapid response to adapt to environmental cues but the mechanism remains unclear.

1.5.7 The role of N7-methylguanosine cap in translation

From yeast to mammals, the N7-methylguanosine cap is required for translation of mRNAs. Removal of the cap structure from *in vitro* transcribed RNA abolishes translation in wheat germ cell-free systems (Muthukrishnan *et al.*, 1975; Shimotohno et al., 1977; Zan-Kowalczewska *et al.*, 1977). When unmethylated capped RNA was added to wheat germ extracts, it was translated only in presence of the methyl donor S-adenosylmethionine (SAM), whereas translation of methylated RNA did not depend on SAM (Both *et al.*, 1975). When *in vitro* methylguanosine capped RNA was microinjected into *X. leavis* oocytes, it was translated more efficiently compared to the guanosine capped RNA (Drummond *et al.*, 1985; Gillian-Daniel *et al.*, 1998). eIF4E was demonstrated to stimulate the translation of specifically methyl capped RNAs but not uncapped RNAs, revealing that eIF4E mediates the role of m7G in translation (Sonenberg et al., 1979; Sonenberg *et al.*, 1980).

In yeast, it has been reported that inactivation of the methyltransferase causes reduction in protein synthesis and the effect is not amenable to reduced transcripts stability (Schwer *et al.*, 2000). Moreover, recent findings suggest that protein translation can be increased when the cap structures physically interacts with the poly (A) tail via PABP. This mechanisms provides that PABP interacts with the poly (A) tail and eIF4G, which in turn stabilises the association of eIF4E with the cap (Kahvejian *et al.*, 2005) and circularises the mRNA (Wells *et al.*, 1998). This interaction results in a close loop that enhances the recycling of ribosomes thus promoting other rounds of translation initiation (Sonenberg & Hinnebusch, 2009; Topisirovic et al., 2011). Although the majority of translation is dependent on eIF4E, a pioneer round of translation was found to be mediated by CBC. This translation does not aim to produce large amount of proteins but functions as mRNA quality control. For example, mRNAs containing a premature termination codon (PTC) are targeted for nonsense mediating degradation (NMD) via CBC. The exchange between CBC pioneers translation and steady state translation is regulated by Importin-α (Sato & Maquat, 2009).

Taken together the evidence discussed above describe how many events within the mRNA life cycle are methyl cap-dependent and therefore highlight the importance of m7G and cap-binding proteins in regulation of gene expression (Figure1.5).

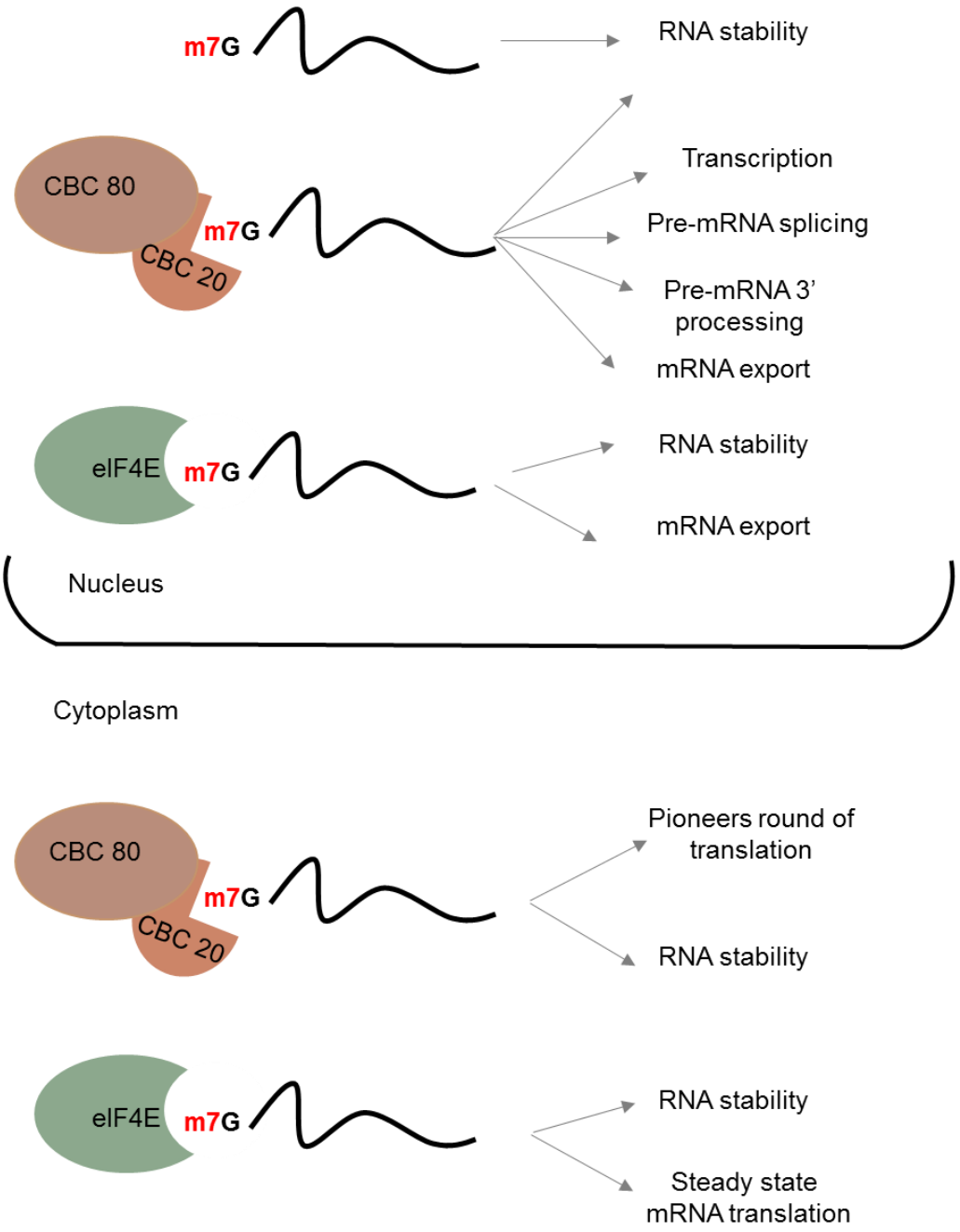


Figure 1.5: Summary of the functions of the N7-methylguanosine cap mediated by the relative by cap-binding proteins.
Most of the functions of N7-methylguanosine cap are mediated by two cap binding proteins: CBC and eIF4E. The former in the nucleus promotes transcription, pre-mRNA splicing and 3' processing and nuclear exports. In the cytoplasm CBC is responsible for a pioneer round of translation. eIF4E promotes the export of mRNAs and in the cytoplasm is required for cap-dependent translation. The methyl cap itself stabilises RNA which is further promoted by the binding of CBC or eIF4E.

1.6 Decapping

The steady state of a given mRNA results from the balance between its transcription and its degradation, therefore an important contribution to regulation of gene expression derives from mRNA decay. In mammals, two main mRNA degradation pathways have been identified and in both mechanisms, bulk mRNA decay initiates with the shortening of poly(A) tail, following which the mRNA body is degraded either from 5'->3' or 3'->5' (Chen & Shyu, 2011). In the 5'->3' decay pathway, the cytoplasmic decapping enzyme Dcp2, which is thought to accumulate in foci, presumably P-bodies, interacts with Dcp1 and hydrolyses the methyl cap of m7GpppG-RNAs longer than 25-29 nt releasing m7GDP (Lykke-Andersen, 2002; Wang *et al.*, 2002). Following the decapping reaction, 5' monophosphorylated mRNAs are targeted for 5' exoribonucleolytic decay by Xrn1. Recently another Nudix hydrolases protein, Nudt16, was found having a decapping activity 5'->3'. It is thought that Dcp2 and Nudt16 preferentially function on different subsets of mRNAs and pathways (Song *et al.*, 2013). More precisely, a comparison of microarray results from $Dcp2^{\beta/\beta}$, β-geo cassette inserted into intron 1 of the Dcp2 gene, Nudt16 knockdown and $Dcp2^{\beta/\beta}$ /sh-Nudt16 MEFs demonstrated that only a subset of mRNAs are jointly regulated by Dcp2 and Nudt16, with their stabilities increasing upon reduction of both proteins (Song et al., 2013; Song *et al.*, 2010). Further studies were based on MEF cells depleted of Dcp2 or Nudt16 and transfected with plasmids encoding either transcripts lacking a nonsense mutation or its nonsense mutation-containing. Comparison between the two transcript populations was analysed by quantitative real-time PCR and authors shown that NMD preferentially utilises Dcp2 over Nudt16, Dcp2 and Nudt16 are

redundant in microRNA-mediated silencing but differentially used for ARE-mRNA decay (Li *et al.*, 2011).

In the 3'->5' decay, following the deadenylation, the RNA is degraded by the exoribonuclease complex, called the exosome (Chlebowski *et al.*, 2013). In this process, the m7G is further hydrolysed by a scavenger decapping enzyme called DcpS, that specifically targets methyl capped RNA 10 nt in length (Liu *et al.*, 2002; Wang & Kiledjian, 2001).

In mammals, DXO decapping enzyme has been recently identified to function as quality control mechanism that targets incomplete cap structure pre-mRNAs (Jiao et al., 2013). It was reported that aberrant capped pre-mRNAs do not proceed into normal RNA processing of splicing and polyadenylation but they are degraded by DXO. All these different complexes give an idea about the specialized mechanism cells evolved to tightly control and regulate the release of viable and functional transcripts and simultaneously prevent the accumulation of potentially deleterious mRNAs.

1.7 Embryonic stem cells

The processes discussed above describe the gene expression pathway and the important role that the synthesis of the methyl cap plays within the different steps of the process. As previously mentioned, through regulation of gene expression, cells spatiotemporally express specific genes required to fulfil their biological role. For example, a specific gene expression machinery allows embryonic stem cells to self-renew while maintaining the concomitant possibility to differentiate into any cell type according to the environmental stimuli (Young, 2011).

1.7.1 What is an embryonic stem cell?

In 1970 it was reported that early mouse embryos grafted into adult mice generated malignant multidifferentiated tumours known as teratocarcinomas (Solter *et al.*, 1970; Stevens, 1970). Within the tumour there are undifferentiated cells, named embryonal carcinoma (EC), which can be propagated in culture and are able to differentiate in all three germ layers: ectoderm, mesoderm and endoderm, like EC cells derived from spontaneous germ cell teratocarcinomas a few years earlier (Kleinsmith & Pierce, 1964; Martin & Evans, 1975). Therefore, the stem cells of teratocarcinomas were the first self-perpetuating pluripotent cells to be characterized. In 1981, permanent pluripotent cell lines were derived directly from the inner cell mass (ICM) of mouse blastocyst. These cells exhibited the same features of teratocarcinomas stem cells and were termed embryonic stem cells (ESCs) to specify the embryo origin and to

distinguish them from embryonal carcinoma (EC) cells (Evans & Kaufman, 1981; Martin, 1981).

ESCs harbour three hallmarks: they undergo symmetrical self-renewing divisions, they are pluripotent as they are able to differentiate into any fetal and adult cell lineages, and if incorporated into blastocyst-stage embryos, they can contribute to functional tissue generation (Smith, 2001). The stem cells potential to generate both stem cells and differentiated cells generated controversy in the field for the type of division stem cells may undergo. In fact, one strategy by which stem cells could accomplish both self-renewal and ability to differentiate is via asymmetric cell division, whereby each stem cell divides to generate one daughter with a stem cell fate and the other able to differentiate (Clevers, 2005). On the other hand, with the symmetrical cell division each stem cell can divide symmetrically to generate either two daughter stem cells or two differentiated cells. Within a population, a pool of stem cells with equivalent developmental potential may produce only stem-cell daughters in some divisions and only differentiated daughters in others (Morrison & Kimble, 2006). Therefore, symmetrical division is required to expand a stem cell pool or to commit stem cell to differentiation. The evidence for symmetrical cell division is well documented in *C. elegans*, Drosophila and vertebrates and are well described in wound heling and regeneration (Morrison & Kimble, 2006). The unlimited expansion and the multilineage differentiation of ESCs cells, caught the worldwide scientific interest and few years later, human ESCs (hESCs) cells were also derived from human blastocyst (Thomson *et al.*, 1998). The unique features of ESCs led to expectations that these cells might be useful to treat a host of degenerative diseases, such as Parkinson's disease and diabetes, as

well as injuries, such as spinal cord injury. However, clinical application of hESCs cells raises issues about the ethical use of human embryos and problems with tissue rejection after implantation. A major breakthrough in the field was the generation of pluripotent cells directly from somatic cells (Takahashi & Yamanaka, 2006). By reprogramming somatic cells, the ethical issues were circumvented. Once established, these cells may be used in regenerative medicine and also to elucidate disease mechanisms and to screen drugs.

1.7.2 How is pluripotency maintained?

1.7.2.1 Extrinsic factors

For each cell division, ESCs have to decide whether to self-renew or to differentiate. The instructions and signals to balance the two choices come from the growth factors present in the external environment and from internal regulators. Mouse ESCs (mESCs), were initially co-cultured with mitotically inactivated mouse embryonic fibroblast (MEFs) (Evans & Kaufman, 1981; Martin, 1981), but later studies on leukemia inhibitory factor (LIF) null fibroblasts elucidated that the fibroblast feeder layer dependency was due to the production of LIF (Smith *et al.*, 1988; Stewart *et al.*, 1992). LIF belongs to the interleukin-6 cytokine family and its effect is mediated by heterodimerization of the cytokine receptors LIF receptor (LIFR) and gp130. Stat3 is then recruited to the receptor complex where it is phosphorylated by Jak, resulting in its subsequent dimerization, nuclear translocation and target gene activation (Darnell, 1997). Inactivation of Stat3 promotes spontaneous differentiation thus

Stat3 is essential for LIF-R/gp130-mediated mESCs self-renewal (Niwa *et al.*, 1998). However, the findings that cells require the presence of serum coupled with the evidence that hESCs do not need LIF, suggested that other external regulators could play a role in the maintenance of pluripotency. In fact, it was reported that a critical signal contribution comes from the bone morphogenetic proteins (BMPs), which act via the Smad pathway and induce expression of inhibitor of differentiation (Id) genes to suppress differentiation (Ying *et al.*, 2003a).

Additionally, a role for Wnt-dependent signalling in self-renewal of human and murine ESCs was reported. The activation of Wnt pathway leads to inhibition of Glycogen Synthase Kinase 3 (GSK3) that in turn results into nuclear accumulation of β-catenin, which eliminates the repressive influence of Tcf3 on the pluripotency network (Wray *et al.*, 2011). Wnt signalling can be activated by direct, intracellular inhibition of GSK3 using specific inhibitor.

Activation of the canonical Wnt pathway was shown to maintain the undifferentiated phenotype as it sustains the expression of pluripotency factors such as Oct4 and Nanog (Sato *et al.*, 2004). Therefore, LIF, BMP and Wnt, together with others, synergistically act to maintain self-renewal and pluripotency of mESCs. Some pathways have also common targets as for example LIF and Wnt that converge on c-Myc, whereby the former ensures the gene transcriptional activation and the latter c-Myc protein stability through inhibition of T58 phosphorylation (Cartwright *et al.*, 2005). It is well established that the culture conditions required for maintenance of mESCs and hESCs are different. hESCs do not need LIF but instead require Fgf2 and Activin/Nodal signalling to sustain self-renewal (Ohtsuka & Dalton, 2008).

1.7.2.2 Intrinsic factors

Cells sense and respond to their cellular and biochemical environment through signal transduction pathways, which can deliver information to the genome in the form of activated transcription factors or cofactors. Transcription factors (TFs) are class of proteins that bind to both promoter proximal and distal DNA elements of target genes and in mammals, approximately 10% of all protein-coding genes encodes for transcription factors emphasising the important role of these proteins (Levine & Tjian, 2003). In ESCs, Oct4, Sox2 and Nanog constitute the “core” pluripotency factors required to maintain stem cells in undifferentiated state (Avilion *et al.*, 2003; Chambers *et al.*, 2003; Mitsui *et al.*, 2003; Nichols *et al.*, 1998; Niwa *et al.*, 2000).

Oct4 belongs to the POU factors which recognise the octamer motif ATGCAAAT. The POU class includes the Pit, Oct and Unc transcription factors that interact with DNA through two DNA-binding domains: POU-specific and a higher affinity POU-homeodomain, connected by a flexible linker (Phillips & Luisi, 2000). In the mouse blastocyst, Oct4 is expressed in the ICM but downregulated in trophectoderm and when *Oct4*$^{-/-}$ embryos were cultured *in vitro*, the developmental potential was restricted to trophectodermal lineage since none of ICM-derived structures were observed (Nichols et al., 1998). Moreover, a study based on Oct4 conditional expression and repression showed that increase in Oct4 levels causes differentiation into primitive endoderm and mesoderm, whereas Oct4 repression results in dedifferentiation to trophectoderm (Niwa et al., 2000).

Sox2 is a member of the Sox (SRY-related HMG box containing) family that binds DNA through the 79-amino acid high mobility group (HMG) domain. Similarly to Oct4, Sox2 is expressed in the ICM, but also in the extra embryonic ectoderm in the early embryo and in developing central nervous system (CNS) (Avilion et al., 2003). As for Oct4, *Sox2*-null cells differentiated primarily into trophectoderm (Masui *et al.*, 2007). Oct4 and Sox2 bind to adjacent sites within the enhancer of target genes resulting in a ternary complex protein-protein-DNA. The physical interaction is a prerequisite for the enhancer activity and it is now established that Oct4 and Sox2 function as heterodimer binding to common target genes regulating their expression (Ambrosetti *et al.*, 1997; Ambrosetti *et al.*, 2000; Botquin *et al.*, 1998; Nishimoto *et al.*, 1999; Williams *et al.*, 2004; Yuan *et al.*, 1995). Moreover, it has been shown that Sox2 contributes to pluripotency by regulating Oct4 levels (Masui et al., 2007).

Nanog protein contains a homeodomain which shares a maximum of only 50% amino acid identity with members of the NK2 family, making Nanog a divergent homeodomain protein. Nanog was identified by two distinct groups either with functional cDNA expression cloning or an *in silico* differential expression analysis (Chambers et al., 2003; Mitsui et al., 2003). Nanog is first detected in the morulae and in the ICM but it is downregulated prior to implantation. *nanog*-deficient ICM failed to generate epiblast as they expressed markers of both parietal and visceral endoderm (Mitsui et al., 2003). On the contrary, constitutive Nanog expression confers cytokine independent self–renewal to undifferentiated ESCs and based on this phenotype the gene was named after Tir Na Nog, the Land-of-the-Ever-Young Celtic myth (Chambers et al., 2003).

ESCs with reduced Nanog levels have higher propensity to differentiate unless Nanog is re-expressed (Chambers *et al.*, 2007). Differently from the uniform Oct4 levels, Nanog is heterogeneously expressed within the cell population as low-Nanog-level cells, with higher propensity to differentiate, coexist with high-Nanog-level cells, which on the contrary, exhibit high self-renewal efficiency. This heterogeneity confers different degrees of responsiveness to differentiation signals in individual cells and allows to maintain a percentage of cells resistant to differentiation (Abranches *et al.*, 2013; Chambers et al., 2007).

c-Myc, despite not being a part of the core pluripotency network, is another fundamental TF required for stem cell pluripotency (Chen *et al.*, 2008; Hu *et al.*, 2009; Kim *et al.*, 2008). c-Myc binds to physiological target genes through a heterodimer complex with basic region helix-loop-helix/leucine zipper protein Max (Dunn & Cowling, 2015). The N-terminal domain includes the transcription activation domain. Additionally to Max, the proteins TFII-I, BRCA1 and MIZ-1 have been implicated in interactions with this region. Dimerization of Myc with MIZ-1 leads to its recruitment to MIZ-1 responsive genes causing histone deacetylation, promoter DNA methylation and consequently repression of transcription (Peukert *et al.*, 1997). Fundamentally important for Myc to promote transformation is its interaction with transformation-transactivation domain-associated protein (TRRAP), which recruits a histone acetyltransferase (HAT) complex that acetylates histone H4, thereby opening chromatin structure and promoting transcription at Myc-bound genes (Nikiforov *et al.*, 2002). Moreover, a biotin labelling technique has recently identified and validated as DNA helicase protein chromodomain as novel Myc interacting partner (Dingar *et al.*, 2015).

Cartwright and colleagues identified c-Myc as a direct target of LIF/Stat3 pathway and therefore it is required for the maintenance of pluripotency. Following LIF withdrawal, c-Myc downregulation occurs via phosphorylation by GSK3β on Threonine-58 leading to its ubiquitin-dependent degradation (Cartwright et al., 2005). c-Myc works as a universal amplifier of gene expression and this explains the plethora of physiological processes it is involved in (Nie et al., 2012).

1.7.3 Induced pluripotent stem cells

The critical role of the above TFs in specifying stem cell identity was further documented when mouse and human embryonic fibroblasts were reprogrammed to induced pluripotent stem (iPS) cells by overexpression of Oct4, Sox2, c-Myc and Klf4 (Takahashi *et al.*, 2007; Takahashi & Yamanaka, 2006). This evidence proved that those factors have a hierarchical role in driving expression of genes that are required to maintain the undifferentiated state. The first reprogramming process was achieved through retroviral transduction with 24 transcription factors highly expressed in ES cells. This cadre of genes was gradually reduced to four that encode the transcription factors octamer Oct4, Sox2, Klf4, and c-Myc (Takahashi & Yamanaka, 2006). A major limitation of reprogramming strategies was the use of potentially harmful genome integrating viruses to deliver reprogramming factor transgenes. Therefore other alternative gene factor delivery systems have been approached including the doxycycline-inducible excisable piggyBac (PB) transposon system (Woltjen *et al.*, 2009). iPS cells exhibit the morphology and growth properties of ES cells and express ES marker genes, however controversy in the field comes

from evidence that question the similarity of the two cell types. In fact, recent works comparing the gene expression profiles of ESCs and iPSCs have suggested that iPSCs, regardless of their origin or the method by which they were generated, are a unique cellular subtype distinct from ESCs (Chin *et al.*, 2009). ChIP followed by hybridization to a human promoter array indicated that the hiPSCs had not acquired a completely novel epigenetic identity and that hiPSCs have a miRNA signature that defines them as unique from hESCs (Chin et al., 2009; Marchetto *et al.*, 2009).

Moreover, despite the generation of iPS cells offers promising opportunities for patient-specific pluripotent cell-based regenerative medicine, a major concern is their potential to develop tumours. A part form Myc, which is a well established oncogene, also the other reprogramming factors are also known to be highly expressed in various types of cancer (Ben-David & Benvenisty, 2011). It was also reported that reactivation of the reprogramming factors predispose iPSCs to genomic instability and thus promotes tumorigenicity (Ramos-Mejia *et al.*, 2010). Similar data were reported for murine iPS whereby histological analysis showed that mice derived from iPSCs developed different tumours in different organs (Tong *et al.*, 2011). Therefore, a more comprehensive knowledge of the reprogramming process is crucial for future clinical applications of iPS cells.

1.8 Gene expression in stem cells

1.8.1 The core pluripotency network

To gain more insight into how pluripotency is maintained, genome wide analyses in mouse and human ESCs were performed (Boyer *et al.*, 2005; Loh

et al., 2006) and three main observations were made. Firstly, the core pluripotency factors Oct4, Sox2 and Nanog co-occupy hundreds of target genes recruiting other TFs to form multiple transcription factor binding loci (MTL) (Chen et al., 2008; Kim et al., 2008; Wang *et al.*, 2006). Secondly, amongst the hundreds of target genes, the core TFs bind both actively expressed genes required for pluripotency but also lineage-specific genes silent in ESCs.

Thirdly, each factor binds its own promoter suggesting the presence of self-reinforcing autoregulatory loop mechanism (Jaenisch & Young, 2008). Oct4 and Sox2 positively regulate each other as they bind to the composite sox-oct elements in both *Pou5f1* (gene encoding for Oct4) and *Sox2* generating a positive feedback loop (Chew *et al.*, 2005). This architecture presumes that the pluripotency factors could be homogenously expressed and initially, also Nanog was included in this positive feedback (Loh et al., 2006). However, contrary to Oct4 levels, Nanog is heterogeneously expressed within a cell population (Chambers et al., 2007). Recent evidence revealed that Nanog represses *Nanog* gene activation and this autorepression contributes to heterogeneous Nanog expression (Fidalgo *et al.*, 2012; Navarro *et al.*, 2012). The ability of Oct4, Sox2 and Nanog to positively regulate genes required for the maintenance of pluripotency, while repressing lineage commitment genes, may explain the ability of ESCs to self-renew and, at the same time, be able to differentiate in response to development cues. To accomplish this, pluripotent cells exhibit a peculiar gene expression landscape different from differentiated cells.

1.8.2 Chromatin structure in stem cells

Undifferentiated cells exhibit a decondensed heterochromatin and a hyperdynamic binding of chromatin-associated proteins (Meshorer *et al.*, 2006). This open chromatin structure confers stem cells the flexibility needed to initiate rapid induction of expression of lineage-specific genes (Efroni *et al.*, 2008; Meshorer et al., 2006). Another peculiar feature of stem cell chromatin is the presence of bivalent domains, which exhibit the concomitant histone modification H3 Lysine 4 trimethylation (H3K4me3) and H3 Lysine 27 trimethylation (H3K27me3), associated with transcription activation and repression, respectively (Bernstein *et al.*, 2006). The H3K27me3 modification is catalysed by Polycomb repressive complex (PRC2) (Lee *et al.*, 2006). During stem cell differentiation many bivalent marks are resolved to a monovalent one, suggesting they maintain the developmental genes in a poised state, ready to be activated when the differentiation pathways are triggered (Mikkelsen *et al.*, 2007). Also chromatin regulatory proteins are actively involved in the maintenance of self-renewal and pluripotency. For example, Tip60-p400 represses differentiation genes to maintain ESCs identity (Fazzio *et al.*, 2008), whereas INO80 complex, selectively occupies promoters of core pluripotency genes and facilitates the recruitment of RNA Pol II to those genes (Wang *et al.*, 2014a).

1.8.3 Control of transcription in stem cells

Genes transcribed by RNA polymerase II typically contain two distinct families of cis-acting transcriptional regulatory DNA elements: (a) a promoter, which is the region of the gene at which RNA Pol II binds to initiate transcription and (b)

distal regulatory elements, which can be enhancers, silencers, insulators, or locus control region (LCR) (Spitz & Furlong, 2012). Enhancers, despite being situated distal to promoters, located upstream or downstream the core promoter, increase gene expression by recruiting different TFs, RNA Pol II and chromatin remodelling enzymes. ChIP-based high-throughput assays have mapped the binding sites of pluripotency-associated TFs (Chen et al., 2008; Kim et al., 2008) and their co-regulators in mESCs. Two different modules were identified: the “Oct4-centric” and “Myc-centric” modules. The core pluripotency network factors and c-Myc are bound at different position across the genome and their location is intimately connected to the mechanism by which they regulate transcription in stem cells. The “Oct4-centric” includes the Oct4, Sox2 and Nanog triumvirate additionally to Zfp281, Esrrb, Nr5a2, Tcfcp2l1, Klf4, Smad1, Stat3 and Tcf3, the last three being the downstream effectors for signalling pathways controlled by BMP, LIF and Wnt respectively. This module clearly elucidates the interplay between the extracellular signals and the core transcriptional regulatory network (Boyer et al., 2005).

The “Oct4-centric” module is mainly located at enhancer sites of their target genes and depending on the association with other factors, it can either promote or repress gene expression. The association of Oct4 with the above TFs as well as with p300 cofactor and the Mediator, which physically and functionally connects the enhancers and the core promoters of active genes, helps to recruit RNA Pol II, thus promoting expression of those genes required for pluripotency (Chen et al., 2008; Kagey *et al.*, 2010; Kim et al., 2008).

Recent studies identified super-enhancers as a cluster of enhancers bound by Oct4, Sox2, Nanog as well as Klf4 and Esrrb (Whyte *et al.*, 2013). Super-

enhancers are mostly found on key pluripotency genes, including Oct4, Sox2 and Nanog themselves. Super-enhancers drive high-level expression of their associated genes (Hnisz *et al.*, 2013) and as a consequence, super-enhancer associated genes are more vulnerable to perturbation of their components and they are often the first genes to be downregulated during ESCs differentiation.

Some developmental genes are marked by the presence of bivalent domains that allow those genes to be silenced to maintain pluripotency but simultaneously poised for activation (Bernstein et al., 2006). It was shown that PCR2 occupies these developmental genes (Ku *et al.*, 2008). The core pluripotency factors were found to occupy the promoters of lineage-specific genes already occupied by PRC2. This co-occupancy reveals an intimate link between repression of developmental regulators and cell pluripotency (Lee et al., 2006). The picture that emerges is that the core pluripotency factors, occupying promoters with different complexes, form a complicated network that drives the expression of pluripotency genes and the repression of lineage commitment ones.

The “c-Myc-centric” module comprises c-Myc, n-Myc, E2f1, Zfx, Rex1 and Ronin (Chen et al., 2008; Kim et al., 2008) and contrary to the core pluripotency factors, this module is found at the promoter sites of paused genes. c-Myc recruits P-TEFb to alleviate RNA Pol II stalling. As mentioned previously, the catalytic subunit of P-TEFb, Cdk9, phosphorylates the CTD of RNA Pol II at Serine-2 position, thus promoting transcription elongation (Rahl et al., 2010).

Therefore, c-Myc rather than acting as a transcription activator, functions as a universal amplifier of all genes that are already “on” but paused (Nie et al., 2012). Combined together, the data suggest that Oct4, Sox2 and Nanog mark

the set of ESCs genes for an active expression and recruit RNA Pol II to those genes, while c-Myc regulates the efficiency of their transcription by promoting RNA Pol II pause release (Young, 2011).

The undifferentiated cells present a unique and complex transcriptional profile and each component of the gene expression machinery collaborates to maintain pluripotency. The general transcription factor IID (TFIID) contains TBP, which recognizes the promoter sequence, additionally to 13 TBP associated-transcription factors (TAFs). The ability of TFIID to bind the promoter makes it essential for the initiation of transcription by RNA Pol II. In mESCs TFIID is highly expressed and its depletion induces downregulation of pluripotency genes and concomitant upregulation of lineage–specific genes. These findings lead to the conclusion that TFIID is required to maintain the transcriptional program of undifferentiated cells (Pijnappel *et al.*, 2013).

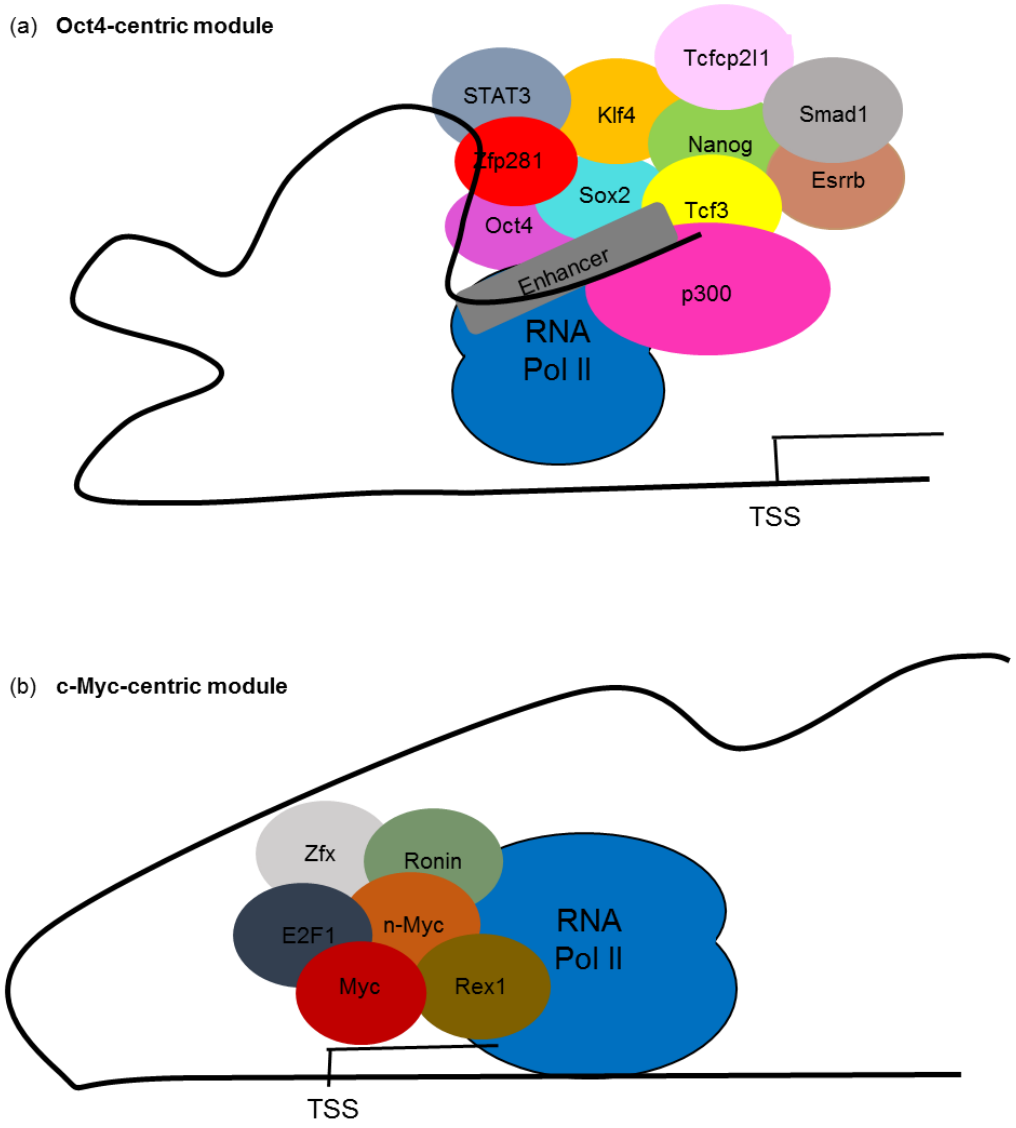


Figure 1.6: Illustration of the Oct4-centric and c-Myc-centric modules.
ChIP-based high-throughput assays mapped the binding sites of pluripotency-associated transcription factors and their co-regulators in mESCs and two modules were identified. (a) In the Oct4-centric, the core pluripotency network binds at the enhancer site of target genes and interacts with many other transcription factors. The transcriptional co-activator p300 is almost exclusively localised to Oct4/Sox2/Nanog targets and helps recruiting RNA pol II as it bridges the interaction between the enhancer-bound transcription factors and RNA Pol II. (b) The c-Myc module is mainly found at the transcription start site and promotes transcription of a large population of genes in ESCs by alleviating the RNA Pol II pause.

1.8.4 Translational regulation in stem cells

The abundance of chromatin-remodelling factors in ESCs endows the genome a preferential open conformation that facilitates the access of the transcriptional machinery and the stochastic formation of PIC even on silenced genes. These features result in a global transcriptional activity, which accounts for another hallmark of undifferentiated state (Efroni et al., 2008). It was also found that in ESCs, normally repressed genome regions, such as satellite repeats, transposons and many tissue-specific genes are stochastically transcribed at low levels, and as soon as cells undergo differentiation, global transcription is gradually reduced, most prominently in these intergenic regions. However, whether the permissive transcripts generate functional proteins was not clear at that time.

Sampath and colleagues started to bridge the gap between transcriptome and proteome profiles that arises during ESCs differentiation. They studied the gene expression between undifferentiated cells and embryoid bodies performing a translation state array analysis (TSAA) that combines global assessment of ribosome loading with microarray analysis (Sampath *et al.*, 2008). They observed that, despite undifferentiated cells present a global transcriptional activity, they display a parsimonious translation, as indicated by the abundance of free ribosomes. On the contrary, when stem cells differentiate, an increase in transcript engagement with ribosomes occurs, leading to a higher translation efficiency, therefore elevating protein synthesis.

The hypothesis of tight translation levels in pluripotent cells is further supported by a recent study about the eIF4E-binding proteins (4E-BPs). Initiation is the rate-limiting step of translation and is subjected to an extensive control (Sonenberg & Hinnebusch, 2009). At this step, mRNA is recruited to the ribosome by eIF4F complex, which amongst other factors, contains the cap-binding protein eIF4E. 4E-BPs are translational inhibitors that when dephosphorylated avidly bind eIF4E inhibiting the assembly of eIF4F complex, consequently preventing translation initiation. mESCs maintain 4E-BP1 hypophosphorylated whereas differentiation induces phosphorylation and thus promotes the translation of lineage-specific mRNAs (Tahmasebi *et al.*, 2014). Authors found that the low 4E-BP1 phosphorylation state is a feature of pluripotent cells, both ESCs and iPSCs.

1.9 RNA binding proteins in pluripotency

The study of protein-nucleic acid interaction initially focused its efforts on DNA binding proteins. However, the biochemical properties of RNA, the complexity of the processes in which RNA is involved, and the documented deregulation of RNA binding proteins (RBPs) with several human disorders (Cooper *et al.*, 2009), brought the attention to the biology of RBPs (Sharp, 2009). The technical development of a cross-linking and immunoprecipitation assay (CLIP) (Ule *et al.*, 2005) and the further improvement of the technology (Darnell, 2010), found that numerous RBPs are involved in the maintenance of pluripotency and in the reprogramming processes (Guallar & Wang, 2014; Kwon *et al.*, 2013).

RBPs play an important role in all successive steps that follow pre-mRNA transcription, including maturation, transport and stabilization, thus directly influencing the protein synthesis. Particularly the “Myc-centric” module relies on RBPs to accomplish its regulatory function (Kwon et al., 2013).

1.9.1 RPBs in mRNA processing: splicing and polyadenylation

As mentioned earlier, the splicing consists of the removal of introns and the following joining of exons together (Braunschweig et al., 2013). The alternative splicing (AS) event represents a real strategy to increase the biological complexity referable to a single locus. Recent studies showed that during stem cells differentiation the repertoire of splice isoforms is gradually reduced (Wu *et al.*, 2010) and therefore it is assumed that splice variants are important for the maintenance of pluripotency. The major evidence is that the pluripotency transcription factors Dnmt3b, Nanog, Sall4, and Oct4 are themselves subjected to AS (Das *et al.*, 2011; Rao *et al.*, 2010; Tsai *et al.*, 2014). Importantly, the different splice versions can fulfil different and even opposite roles as shown for Forkhead box (Fox). The FOXP1-ES specific isoform stimulates the expression of Oct4 and Nanog and thus sustains pluripotency, whereas the other variant promotes differentiation (Gabut *et al.*, 2011). SON and Tip110 are RBPs identified in human ESCs, required to specifically regulate the OCT4A, active splicing version of OCT4 and other pluripotency regulators (Liu *et al.*, 2012; Liu *et al.*, 2013; Lu *et al.*, 2013). It has been shown that knockdown of SON results in human ESCs differentiation. In mESCs, Nanog, but not Oct4 or Sox2, was found to be bound to another RBP known as Rbm47 but further studies are needed to elucidate its function (Yeganeh *et al.*, 2013).

Collectively, these findings suggest that the RBPs regulate the AS of pluripotency transcripts in a gene specific manner.

Additionally to methyl cap synthesis and splicing, the third step of mRNA maturation consists in the addition of the poly(A) tail catalysed by a polyadenylate polymerase (PAP) and the cleavage by polyadenylation specificity factor (CPSF). An alternative polyadenylation (APA) generates different 3'UTRs that in turn affect stability, translation and localization of mRNAs and thus their expression (Di Giammartino et al., 2011). It was reported that stem cells are characterized by short 3' UTR (Mueller *et al.*, 2013) and a recent study revealed that Fip1, RBP of the CPSF complex, promotes self-renewal by regulating the APA profile of critical pluripotency markers. In fact, depletion of Fip1 caused a decrease in levels of pluripotency markers and a reduction in the reprogramming efficiency (Lackford *et al.*, 2014). Therefore, APA regulation serves as fine-tuning mechanism for gene regulation in ESCs, adding a new layer of control of self-renewal and pluripotency.

1.9.2 RPBs involved in mRNA nuclear export

Apart from gene specific mRNA processing, a selective nuclear transport constitutes another layer of RNA-related regulation of pluripotency. As previously mentioned, the export of mRNA from the nucleus to the cytoplasm is mediated by several adaptors amongst which is THO/TREX complex. The THO/TREX couples mRNA biogenesis with the nuclear export of mature transcripts. The RPBs Thoc5 and Thoc2, components of the THO complex, were recently identified in genome-wide siRNA screens searching for ESCs

self-renewal factors (Ding *et al.*, 2009). Further studies revealed that their expression is correlated with the pluripotent state and their depletion leads to ESCs differentiation (Wang *et al.*, 2013). In fact, knockdown of the two components caused reduction in the association of polysome with Nanog, Sox2, Klf4 and Esrrb mRNAs resulting in a decrease of their protein levels. However, the impairment in the protein levels was not accompanied by a decrease in their transcript abundance and this discrepancy was caused by retention of the respective transcripts in the nucleus, thus preventing them from translation. Therefore it was concluded that THO properly directs some pluripotency transcripts for nuclear export to the cytoplasmic translation machinery in a Thoc5 dependent manner. Interestingly, Oct4 expression was not affected by Thoc2 or Thoc5, further supporting that RBPs exert their functional role on some pluripotency transcripts but not others, thus regulating pluripotency in a gene specific manner.

Adaptor proteins interact with the nuclear receptors to mediate the export of mRNAs, whereas karyopherins, including both Importin and Exportin, are mobile targeting receptors that mediate the bidirectional trafficking of macromolecules across the nuclear envelope (Fried & Kutay, 2003). Also karyopherins collaborate to maintain the undifferentiated cell state, by modulating the import/export of specific proteins. Importin-α binds the cargo through its NLS binding site. Following the recognition of the cargo NLS, Importin-α dimerizes with Importin-β which directly interacts with the nucleoporins allowing the cargo translocation. There are several Importin-α family members and a recent study revealed that a particular subtype, Importin-α2, is required to maintain pluripotency.

The mechanism proposed is that in undifferentiated cells, Importin-α2 is able to selectively mediate Oct4 nuclear import, while retaining differentiation factors Oct6 and Brn2 in the cytoplasm (Yasuhara *et al.*, 2013). Importin-α2 is highly expressed in ESCs and its expression is reduced during differentiation, whereas Importin-α1/4 levels exhibit the opposite trend, possibly to allow the import of the TFs required for the lineage commitment. Therefore, the dual Importin-α2 activity and the various expression patterns of its different subtypes contribute to maintain pluripotency via regulating TFs localization.

1.10 Decapping in stem cells

As discussed earlier, RNA degradation plays an important role in the regulation of gene expression, therefore it is not surprising that mRNA decapping process collaborate to maintain the pluripotency. Utf1, one of the downstream targets of Oct4 and Sox2, apart from preventing the excessive loading of PRC2 on bivalent gene, is also able to repress genes through mRNA pruning (Jia *et al.*, 2012). More precisely it was shown that Utf1 recruits Dcp1a (a non-catalytic subunit of the mRNA decapping complex) to the bivalent promoters and therefore it promotes the degradation of mRNAs transcribed from leaky bivalent genes, actively participating in the repression of genes that are not required for pluripotency.

1.11 mRNA modification

The RNA modifications join the long list of multiple layers of gene expression regulation. The most abundant internal modification of mRNA of all higher

eukaryotes is the N6-methyladenosine (m^6A). A recent research found that in mESCs the two methyltransferase like enzymes, Mettl3 and Mettl14, synergistically work to methylate the RNA of target genes (Wang *et al.*, 2014b). The depletion of the two enzymes by RNAi led to downregulation of the pluripotency markers and upregulation of developmental genes (more enriched in m^6A) suggesting that m^6A is involved in the maintenance of pluripotency. Further analysis demonstrated that the modification prevents the binding of a RNA stabilizer protein HuR. Therefore the model proposed is that the modification, destabilising the transcripts of lineage commitment genes, favours mESCs undifferentiated state.

A parallel study based on Metll3 Knockout (KO) showed opposite results as rather than affecting mESCs viability and self-renewal, the Metll3 KO mESCs renewed at an improved rate and resulted in blocked differentiation (Batista *et al.*, 2014). Consistently with Batista et al work, Guela and colleagues recently observed that depletion of m^6A in mRNA of Metll3 KO mESCs hampers their priming and differentiation competence, which leads to a “hyper”-naïve pluripotency phenotype. Authors conclude that m^6A acts as a timely maintainer of the balance between pluripotency and lineage priming factors, thus ensuring orderly differentiation of mESCs (Geula *et al.*, 2015).

These controversial findings about the effect of m^6A suggest that more studies are required to fully understand how the mRNA modifications affect cell pluripotency.

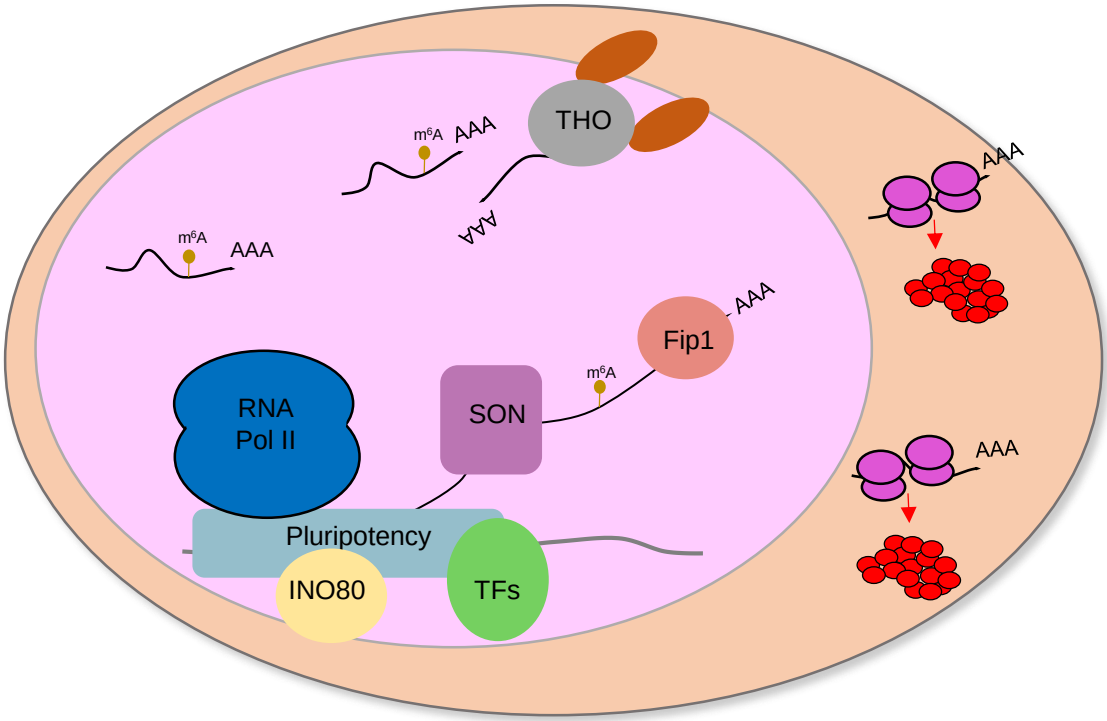


Figure 1.7: Schematic representation of transcriptional and post-transcriptional regulation of the ESCs state.
In order to maintain pluripotency the gene expression program of stem cells is different from differentiated one. Transcription factors, chromatin remodelling enzymes, TFIID of RNA Pol II, alternative splicing, alternative polyadenylation, m^6A and mRNA export function in concert to maintain pluripotency**.**

1.12 Summary and aim of the project

Regulation of gene expression allows cells to spatiotemporally control their transcript and protein content to fulfil their biological role and adapt to the environmental cues. The synthesis of the N7-methylguanosine cap is fundamental within the gene expression pathway as CBC and eIF4E proteins recognising the mRNA cap moiety, promotes RNA Pol II transcription initiation and elongation, mRNA splicing, 3' end pre-mRNA processing, mRNA nuclear export, translation and stability. Recent studies reported that some transcripts are incompletely capped as they lack the methyl moiety on their cap structure and the decay of these aberrantly capped transcripts is mediated by specific decapping machinery (Jiao et al., 2013). Moreover, it was recently reported that growth factors and TFs regulate the formation of the methyl cap. Altogether these evidence strongly support that the m7G rather than being a constitutive process is subjected to an active regulation (Aregger & Cowling, 2013; Chang et al., 2012; Cole & Cowling, 2009). Within the mRNA cap synthesis, RNMT has been found to be associated with a small protein called RAM, which activates the enzyme and contains an RNA binding domain. RAM was recently identified (Gonatopoulos-Pournatzis et al., 2011) and despite its biochemical function being characterised (Gonatopoulos-Pournatzis & Cowling, 2014), its biological relevance remains elusive.

As previously described, embryonic stem cells, in order to maintain the unique features of self-renewal and pluripotency, exhibit a peculiar gene expression landscape, whereby each singular event in the mRNA life cycle is regulated, to maintain the undifferentiated state.

All the events, orchestrated by the main TFs, function in concert to sustain ESCs state by activating pluripotency genes and simultaneously repressing lineage-specific genes (Young, 2011). Based on the important role of RNA binding proteins in the maintenance of pluripotency and considering mRNA cap-dependent events, which are regulating gene expression, the investigation of mRNA cap methylation in stem cells may provide a novel insight into the biology of ESCs. Therefore, this PhD project aims to investigate whether RNMT and RAM are regulated during cell differentiation to ultimately uncover the role of cap methylation in undifferentiated cells.

2 Materials and Methods

2.1 Materials

2.1.1 Lab equipment and consumables

Automatic film processor (Konica Corporation)c, Autoradiography cassette (Siemens), C1000 thermal cycler (Bior-Rad), Cell countess (Invitrogen), Cell Lifter (Corning Incorporated), Centrifuge 5415-R (Eppendorf), CO_2 incubators (Mackay+Lynn), Eppendorf Tubes (Star Lab), Eppendorf Tubes (Star Lab), Falcon Tubes (Greiner), Falcon Tubes (Greiner), Filter 45 μm (Sartorius), Filter Units (Thermo Scientific), Gel Loading Tips (Star Lab), Gel Loading Tips (Star Lab), Pipette Tips (Star Lab), Phosphor screen (FujiFilm), Phosphorimager FLA-500 (FujiFilm), Pipette Tips (Star Lab), Pipettes (Greiner), Syringe needle (Terumo), Syringe needles (Terumo), Tissue culture safety cabinets (Medical Air technology), Versa Max microplate reader (Molecular Devices).

2.1.2 Chemicals

2-Mercaptoethanol (Sigma), Bridge (Merk), Dithiothreitol (DTT) (Formedium), Ethanol (VWR International), Glycerol (VWR International), Glycine (VWR International), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes) (Sigma), Formaldehyde (Sigma), Hydrochloric Acid (VWR International), Methanol (VWR International), Sodium Chloride (VWR International), Sodium Dodecyl Sulfate (SDS) (Sigma), Sodium Fluoride (Sigma), Tris-Base (VWR International), TritonX-100 (VWR International), Tween 20 (VWR International), Sucrose (VWR International).

2.1.3 Mammalian cell culture and maintenance

100 mm Petri Dishes (Helena), 150 mm Petri Dishes (Helena), 6-Well Plate (Greiner), A8301 (DSTT), B27 supplement serum free (Gibco), Cover Glass 13 mm (VWR International), Cryovials (Alpha Labs), DAPI (Sigma), DPBS (Invitrogen), DPBS (Sigma), Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen), Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (Gibco), Fetal Bovine Serum (FBS), .C. Approved (Invitrogen), Gelatin from porcine skin (Sigma), Glasgow MediGlasgow Minimum Essential Medium (GMEM) (Sigma), Kanamycin (Sigma) Knockout Serum Replacement (Gibco), L-Ascorbic acid (Sigma), L-Glutamine (Gibco), Lipofectamine 2000 (Invitrogen), MG132 (Sigma), Lipofectamine RNAiMax (Gibco), Neurobasal media (Gibco), Non-Essential Amino Acids (Gibco), Penicillin/Streptomycin (Invitrogen), Puromycin (Sigma), Sodium Pyruvate (Gibco), Trypan Blue solution (0.4%) (Sigma), Trypsin-EDTA solution (Invitrogen).

2.1.4 Protein analysis

40% Acrylamide : Bis-acrylamide 29:1 (Flowgen Biosciences), Ammonium Persulfate (APS) (Sigma), Aprotinin (Sigma), BenchMark Prestained Protein Ladder (Invitrogen), Bovine Serum Albumin (BSA) (NE BioLabs), DAPI Fluorescent Stain (Invitrogen), Donkey Serum (Sigma), Dried Skimmed Milk (Marvel), Immobilon Transfer Membrane (Millipore), Leupeptin Hydrochloride (Sigma), NNN'N'-Tetramethyletylenediamine (TEMED) (VWR International), Pepstatin (Sigma), Phosphatase Inhibitor Cocktail 2 (Sigma), Phosphatase Inhibitor Cocktail 3 (Sigma), Protein A/G PLUS-Agarose (Santa Cruz Biotechnology), X-RAY FILM (Konica Minolta).

Table 2.1: Buffers and solutions employed for experiments in this study.

Buffer	Composition
F-Buffer	10 mM Tris (pH 7.05), 50 mM NaCl, 30 mM Na pyrophosphate, 50 mM NaF, 10% Glycerol, 0.5% Triton
SDS Running buffer	25 mM Tris, 0.1% SDS, 250 mM glycine
Transfer buffer	25 mM Tris, 192 mM glycine, 20% methanol
TBST	25 mM Tris (pH 8.1), 155 mM NaCl, 0.1% tween 20
4x Laemmli buffer	242 mM, 10% SDS, 25% glycerol, 100 mM DTT, bromophenol blue
MT assay buffer	50 mM Tris (pH 8.0), 6 mM KCl, 1.25 mM $MgCl_2$
Ponceau	Ponceau 5 % (v/v) Acetic Acid, 0.1 % (w/v) Ponceau S
Gel filtration calibration buffer	25 mM HEPES, 10 mM NaCl, 0.03% Bridge, 1 mM DTT
Polysome Lysis buffer	10% Triton x-100, 10 % sucrose buffer, 50µL RNAsin

2.2 Methods

2.2.1 Transformation of *E.coli* and plasmid purification

For each transformation 25 µL of competent *E. coli* DH5α were incubated with 2 µg of plasmid DNA on ice for 20 min. To facilitate the incorporation of the plasmid, heat-shock was performed by incubating cells at 45°C for 45 sec following by 5 min incubation on ice. Cells were then streaked on selective LB agar plates containing the appropriate antibiotic (50 µg/mL ampicillin, 50 µg/mL kanamycin) and inverted plated were incubated at 37°C overnight. A single isolated colony was then picked and grown in 200 mL of LB at 37°C overnight. In case of kanamycin selection the bacteria were incubated in SOC media for 1hr at 37°C before plating.

Cells were then collected by centrifugation at 4000 rpm for 10 min at 4°C. DNA QIAprep Miniprep kit (Qiagen) was used to purify 5-10 µg DNA eluted in 50 µL distilled water according to manufacturer's instructions. To produce large amount of plasmids DNA Maxi kit (Qiagen) was used according to manufacturer's instructions. Typically 500-1000 µg DNA were resuspendend in 1 mL of distilled water.

2.2.2 Cloning

Sequencing of the constructs was performed by the DNA sequencing service (College of Life Science, University of Dundee). All the constructs employed in this thesis (Table 2.2) encode the human orthologous of the gene and were cloned by Dr. Mark Peggie from the Division of Signal Transduction Therapy (DSTT) cloning team. All the constructs were used to transfect mESCs to make stable cell lines expressing the protein encoded in the plasmid.

Table 2.2: List of cDNA constructs utilised for the purpose of this study.

Construct	Vector	Clone Number
Empty Vector	pPyPCAGGS	DU43266
RAM-GFP	pPYPCAGGS RAM GFP WBL	DU42605
FLAG-RAM	pPYPCAGGS FLAG-RAM	DU 42391
FLAG-RAM S36	pPYPCAGGS FLAG-RAM S36	DU 42567

2.2.3 DNA concentration and determination

Nanodrop Spectrophotometer 1000 (Thermo Scientific) was employed to determine the concentration of purified DNA. A baseline measurement was performed using distilled water. The absorbance was measured at 260/280 nm according to manufacturer's instructions.

2.2.4 Cell culture and maintenance

All cell culture was carried out in a class II hood, using aseptic techniques, sterile equipment and reagents. All cell lines were cultured in at 37°C under humidified atmosphere containing 5% CO_2. The 46C mouse embryonic stem cells were kindly provided by Dr. Marios P Stavridis (University of Dundee). mESCs were cultured in GMEM (Sigma) supplemented with 10% Knockout Serum Replacement (Invitrogen), 1mM Sodium pyruvate (Gibco), 1% Non-Essential Amminoacids (Gibco), 0.35 µM beta-Mercaptoethanol (Gibco) and 100 Unit of leukemia inhibitory factor (LIF). This media composition will be further referred as mESCs media.

Primary MEFs were derived from 12.5 days old 129 Ola mice strain whereas immortalised MEFs were a kind gift form Ganley's lab. Both primary and immortalised MEFs were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-Glutamine, 100 units/ml penicillin and 0.1 mg/ml streptomycin.

For cell passaging, each flask of mESCs or MEFs was washed twice with Dulbecco's phosphate buffered saline (DPBS), then 1 mL of 0.05% Trypsin-EDTA (Gibco) was added and the flask was put back in the incubator to allow cell detachment. Following trypsinization, fresh media was added to the dissociated cells to neutralise the Trypsin-EDTA, mESCs and MEFs were spun down at 300 g for 5 min and at 1200 rpm for 3 min respectively. Cell pellet was resuspended in fresh media and plated in a new flask. For passaging or seeding mESCs, plates were coated with 0.1% gelatin for 10 min unless differently specified.

2.2.5 Cell counting

Cell suspension was mixed 1:1 with Trypan blue (0.4%) and cells were then counted using a Countess cell counter (Life technologies) according to manufacturer's instructions.

2.2.6 Freezing cells

Cryo-freezing was used for storage of all cell lines. mESCs and MEFs were grown in a 75 cm^2 flask until 80% confluent and were then trypsinised as described above and pelleted at 300 g for 5 min and 1200 rpm respectively. The resultant pellet was resupsended in GMEM media (mESCs) or serum (MEFs) supplemented with 10% dimethyl sulphoxide (DMSO, Sigma) and transferred into a cryo-vial (1 mL aliquots for 1.5 mL cryo-vial). Cells were frozen gradually submerging the cryo-vials in propanol and placing the propanol container (Mr. Frosty) in -80°C. Cells were transferred in liquid nitrogen storage the following day. For cell recovery after liquid nitrogen storage, the cryo-vials were quickly thawed at 37°C and cells were washed in culture media to remove DMSO. Cells were then cultured in a 25 cm^2 flask in the relative appropriate media.

2.2.7 *In vitro* neural differentiation

mESCs were grown in a 75 cm^2 flask until 80% confluent and were then detached as aforementioned. Cell pellet was resupended in N2B27 medium consisting of one volume of DMEM/F12 combined with one volume of Neurobasal media. The mixed media was supplemented with 0.5% N2 component (homemade) (Table 2.3), 2 mM L-glutamine, 1% B27 Supplement,

50 mM beta-mercaptoethanol (all from Gibco). $2.5x10^5$ were seeded on 10 cm plates previously coated for 2 hr with 0.1% gelatin. Media was changed every two days in all experiments unless otherwise indicated.

Table 2.3: List of components used to make N2 supplement.
Final concentration of each component is shown.

Component	Final concentration
Insulin from bovine pancreas (Sigma)	25 µg/ mL
apo-Transferrin human	100 µg/ mL
Progesterone (Sigma)	6 µg/ mL
Putrescine (Sigma)	16 µg/ mL
Sodium selenite (Sigma)	30 nM
BSA Fraction V (7.5%) (Gibco)	50 µg/mL

2.2.8 Generation of stable cell lines

To make mESCs stably expressing the protein encoded in the plasmid (listed in Table 2.2), $1x\ 10^6$ cells were seeded into a well of 6 well dish coated with 0.1% gelatin. Cells were then transfected with 3 µg of DNA using 3 µL Lipofectamine 2000 (Life technologies) according to manufacturer's instructions. 24 hr post transfection, transfected cells were dissociated using 0.05% Trypsin-EDTA and $5x10^4$ cells were seeded into a 10 cm plate previously coated with 0.1% gelatin. Successfully transfected cells were selected using 1.5 µg/mL of Puromicyn.

2.2.9 Reprogramming

Dr. Marios Stavridis kindly donated plasmids used to transfect MEFs: PB-TAP IRI 2OKSMimO (containing the reprogramming factors), AG-rtTA (reverse tetracycline transactivator) and HyPBase (transposase), the latter was kindly

provided by Dr. Keisuke Kaji (Woltjen et al., 2009). MEFs were seeded into a dish of a 6 well dish in mESCs media supplemented with 1 µg/µL Fgf2. The next day, 1.5 µg of DNA (0.5 µg of each plasmid) were combined with 6 µL of Fugene HD and 100 µL of Optimem and used to transfect seeded MEFs. After 24 hr of transfection, media was changed with mESCs media supplemented with 1 µg/mL doxycycline (Sigma), 10 µg/mL Vitamin C (Sigma) and 500 nM A83-01 (Alk5 inhibitor, DSTT) to select transfected cells. Selection was carried out for 7 days following which selection media was substituted with mESCs media. iPSC colonies were isolated from untransfected MEFs and used for analysis.

2.2.10 siRNA transfection

siRNA-mediated RNA-Interference (RNAi) was employed to deplete RNMT and RAM expression level throughout this thesis. mESCs were transfected with siRNAs using Lipofectamine RNAiMax (Gibco) transfection reagent according to manufacturer's instructions. In a well of 6 well dish, $1x10^5$ cells were seeded and immediately transfected with 200pmol of siRNA (Table 2.4) combined with 4 µL of Lipofectamine RNAiMax and 500 µL of GMEM serum free media. For the polysome profile analysis, the siRNA treatment was performed in a 10 cm dish where $2.5x10^6$ cells were seeded and immediately transfected with 800 pmol of siRNA combined with 16µL of Lipofectamine RNAiMax and 800 µL of GMEM serum free media. Cells were left to grow for 48 hr before harvesting.

Table 2.4: Oligonucleotide sequence of the siRNAs used in this study.
All siRNAs employed were obtained from Dharmacon.

siRNA target	Catalogue details	Sequence
scRNA 4	D-001210-04-20	AUGAACGUGAAUUGCUCAA
RNMT 1	D-059345-01-0050	GGAAAGAGGUGAUGUGUCG
RNMT 2	D-059345-02-0050	CGGAUACGAUGCUUAGAAA
RNMT 3	D-059345-03-0050	GCUUAACCCUGGCGGCUAU
RAM 1	D-049592-01-0050	CAAGACAACAGACAAUUUA
RAM 2	D-049592-02-0050	GACCAGAGCCCUACUAUCA
RAM 3	D-049592-03-0050	GCAGAUUCACGAAAGAUGA

2.2.11 Alkaline phosphatase

To assess the pluripotency of mESCs, alkaline phosphatase activity was revealed using Alkaline Phosphatase (AP) staining kit (Sigma) according to manufacturer's instructions. Cells were seeded into a dish of a 6 well dish and when 80% confluent were fixed with fixative solution for 30 sec. Cells were then rinsed in deionised water and incubated in the dark with the staining solution for 15 min. Cells were then washed and stained colonies counted. To reveal AP activity following siRNA treatment, cells were transfected with siRNA as previously described. 48 hr post transfection, cells were dissociated using 0.05% Trypsin-EDTA and ~100 cells per well were seeded into 6-well dish previously coated with 0.1% gelatin. Cells were grown for 7 days and then stained for AP as described above.

2.3 Mammalian Protein analysis

2.3.1 Cell lysis

Cell lysis was performed at 4°C to avoid protein degradation. Culture medium was removed and cells were washed twice with ice-cold PBS. Cells were lysed with F buffer (Table 2.1) supplemented with 1 mM DTT, 0.1 TIU (trypsin inhibitor

unit) Aprotinin, 1 µM pepstatin and 10 µM leupeptin. For immunoblot analysis of phosphorylated proteins the F-buffer was further supplemented with Phosphatase inhibitors (cocktail mixture 2+3) to avoid dephosphorylation of proteins. Cells were collected by scraping and left on ice for 10 min to be then centrifuged at 13000 rpm for 10 min at 4°C to separate cell extracts from cellular debris (pellet). Supernatant was collected and transferred into a new tube. Proteins concentration was estimated using Bradford method.

2.3.2 Protein concentration determination

Protein concentration was determined using Bradford reagent diluted 1 in 5 with distilled water. To generate a standard curve, dilution of Bovine Serum Albumin (BSA) were made in F buffer and 4 µL of each dilution, or F-buffer only used as blank, were added to 200 µL of 1x Bio-Rad protein assay reagent in a 96-well plate. The absorbance was measured at 595 nm with a plate reader. The measurement obtained from the blank was subtracted from all other samples measurements. A standard curve was derived by plotting the absorbance reading of the protein versus the concentration of BSA and a linear equation determined. The absorbance readings of whole extracts were performed in duplicate and the protein concentration of each sample was determined by plotting the average reading on the standard curve.

2.3.3 Immunoprecipitation

All the steps were performed at 4°C to avoid protein degradation. 1 mg of cell extracts were pre-cleared with Protein A/G Plus-Agarose (Santa Cruz) for 30 min on a rotating wheel. The agarose beads were then pelleted at 4000 rpm for 1 min.

The pre-cleared supernatant was incubated with 1 µg antibody (polyclonal anti-RNMT or sheep polyclonal anti-RAM) and with Protein A/G-Agarose on a rotating wheel for 2 hr to precipitate the antibodies. Immunoprecipitates were pelleted by centrifugation at 4000 rpm for 4 min and washed three times with 1 mL of F-buffer. Immunoprecipitated proteins were eluted in 50 µL of 1X Loading dye and 10-15 µL resolved by SDS-PAGE. Loading dye was also added to the input and to the flow through which were analysed alongside with the IPs.

2.3.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed to resolve proteins according to the molecular weight. The resolving gel composition was made of 400 mM Tris-HCl pH 8.8, 0.1% SDS and 8-15% acrylamide and water. The stacking gel was made combining 400 mM Tris base pH 6.8, 0.1% SDS, 5% acrylamide and water. To trigger the polymerization reaction APS and TEMED were added to final concentration of 0.1% (v/v) and 0.01% (v/v) respectively. Homemade gels were resolved in a Mini-PTROEAN Tetra Electrophoresis system (Bio-Rad) using SDS Running buffer (Table 2.1). Protein samples were combined with 1x Laemmli buffer (Table 2.1) and denatured at 100°C for 5 min and then centrifuged ad 13000 rpm for 1 min. In the first lane of each gel 5 µL of molecular weight markers were loaded. Marker and protein samples (Gibco) were separated at 150 V. The amount of protein resolved and the percentage of acrylamide gel used in the resolving gels for the detection of specific proteins are shown in table 2.5.

Table 2.5: List of the protein resolved by SDS-PAGE in this study.
The molecular weight of the proteins, the amount of total protein extracts analysed and the percentage of acrylamide gel are indicated.

Protein analysed	Molecular weight (kDa)	Amount of protein analysed (µg)	% of Acrylamide gel
RAM	15	15	15
RAM S36	15	15	15
RNMT	64	8	8
Oct4	54	8	8
Sox2	35	15	15
Klf4	54	10	15
Nanog	35	10	15
Actin	50	8	8
E-cadherin	120	10	8
Vimentin	54	15	8
c-Myc	70	15	10
GFP	27	15	15
Pax6	47	15	8
Flag (M2)	9	15	15

2.3.5 Western blot analysis

After separation by SDS-PAGE, resolved proteins were transferred onto polyvenylfluoride (PVDF) membranes in transfer buffer (Table 2.1) using Mini Trans-Blot Cell (Bio-Rad) at 66 V for 90 min at 4°C. The transferred proteins were visualised by Ponceau S solution (Table 2.1) to ensure the efficient transfer.

Subsequently, membranes were blocked for 60 min with TBST (Table 2.1) containing 5% milk or 3% BSA according to the primary antibody. Incubation with primary antibody was carried out overnight at 4°C. List of primary antibodies, blocking solution, working dilution and species where the secondary antibody was raised are listed in Table 2.6. Following incubation with the primary antibodies membranes were then washed three times with TBST to

remove eventual unbound primary antibody. The membranes were then incubated for 45 min with the appropriate secondary antibody at room temperatures. The membranes were then washed five times with TBST. Immunoreactive proteins were visualised on X-ray film using enhanced chemiluminescence reagent according to manufacturer's instructions.

Table 2.6: List of antibodies used for Western blot.
The dilution of the antibody and the blocking reagent used are indicated**.**

Target	Company	Cat N	WB con	Blocking	2nd Ab Dilution	Species
RAM	In house	/	1:100	BSA	1/1000	Sheep
RAM S36	In house	/	1:1000	BSA	1/1000	Sheep
RNMT	In house	/	1:1000	Milk	1/5000	Sheep
Oct4	Abcam	18976	1:1000	BSA	1/5000	Rabbit
Sox2	Cell Signalling	4900	1:1000	5% BSA	1/1000	Mouse
Klf4	Santa Cruz	20691	1:1000	BSA	1:2000	Rabbit
Nanog	Abcam	80892	1:1000	BSA	1:2000	Rabbit
Actin	Abcam	3280	1:1000	Milk	1:5000	Mouse
E cadherin	BD	610181	1:5000	Milk	1:5000	Mouse
Vimentin	Abcam	92547	1:1000	BSA	1:4000	Rabbit
c-Myc	Cell Signalling	9402S	1:1000	BSA	1:2000	Rabbit
GFP	Roche	11814460001	1:1000	BSA	1:1000	Rat
Pax6	Santa Cruz	11357	1:1000	BSA	1:1000	Rabbit
FG (M2)	Sigma	F1804	1:1000	BSA	1:1000	Mouse

2.3.6 Gel filtration analysis

mESCs were collected and lysed as previously described and passed through a 20G needle 10 times and centrifuged at 14,000 g for 10 min at 4ºC. The supernatant was centrifuged at 100,000 g for a further 60 min at 4ºC. 1 mg cell extract was resolved on a Superose 6 column (10x300mm, 13µm particles, 40nm pores, GE Healthcare). Before the run, the column was equilibrated in calibration buffer (Table 2.1). 0.5 ml fractions were collected, combined with Laemmli buffer and analysed by SDS-PAGE followed by western blot to identify elution profile of our proteins of interest

2.3.7 Immunofluorescence

Cells were seeded into round cover slides previously washed with Ethanol 70%. For the *in vitro* neural differentiation, cover slides were coated with Laminin (Sigma L2020) for 3 hrs. The whole protocol was performed at room temperature. The day of collection, culture media was removed and cells were washed twice with ice-cold PBS to be then fixed with fresh 4% paraformaldehyde for 30 min. Cells were then washed in PBS and permeabilised in 1% Triton X-100 diluted in PBS. Following which, cells were blocked with 1% Donkey serum in PBS with 0.05% Tween 20 (primary and secondary antibody dilution solution) for 30 min. Primary antibodies (Table 2.7) were incubated for 1 hr in antibody dilution solution. Cells were then washed three times with PBS to be subsequently incubated with the appropriate secondary antibodies (Alexa) diluted 1:500 in dilution solution in the dark for 1 hr. Cells were washed three times with PBS to remove any unbound antibody and stained with DAPI 1:15000 in antibody dilution solution for 5 min.

Finally, cells were washed with distilled water and mounted on glass plates. High-resolution images were collected with an imaging system (DeltaVision Restoration; Applied Precision) using a 40X/ 1.514 oil (Olympus) objective lens. Images were then processed using OMERO software unless otherwise specified.

Table 2.7: List of antibodies used for immunofluorescence.
The species where antibodies were raised and the dilution used for this study are shown.

Antibody	Company	Working dilution
RNMT	In house	1:500
RAM	In house	1:100
Oct4	Abcam 18976	1:500
Pax6	Santa Cruz 11357	1:500
Tubulin-βIII	Abcam 11309	1:500

2.4 RNA Extraction and Analysis

2.4.1 RNA extraction

The culture medium was removed and cells were washed twice with ice-cold PBS. RNA was then isolated with RNeasy GeneJet RNA purification kit (Thermo Scientific) following manufacturer's instructions and quantified using Nanodrop.

2.4.2 Quantitative reverse transcriptase (RT-qPCR)

500 ng of RNA was converted into complementary DNA (cDNA) with random hexamer primers using the iScript cDNA Synthesis Kit (Quanta) according to

manufacturer's instructions. The converted cDNA was diluted 1:10 in distilled water and 3 µL were subjected to qPCR with a Bio-Rad iQ5 RT-PCR detection system. The total volume of qPCR reaction was 10 µL using 5 µL of Quanta Bioscience SYBR Green FastMix for iQ and 0.3 µM of appropriate primers (Table 2.8). Each sample was loaded in duplicate to avoid pipetting errors. The Critical threshold (Ct) values were determined and expression values were normalised against the reference gene, e.g. Actin.

Following siRNA treatment, the expression levels in RT-qPCR assay were obtained using ΔΔC(t) of the average of duplicate samples and the average of loading control, e.g. Actin.

Table 2.8: List of primers employed for RT-qPCR study.
Sequence and melting temperature of primers used in this study are shown. The species specificity of a primer pair is mentioned next to its name, e.g. (m) for mouse and (h) for human primers.

Locus name (Species)		**Primers**	**T (a)**	**N**
RNMT (m)	F	GCAGGCGGATACGATGCTTA	59°C	20
	R	CTTTGCCATTATCAGTTCAAAGCTA		25
RAM (m)	F	AGTCGGGTGGTTGAGGGATT	60°C	20
	R	TTAAGATCCTTCATGGGACGCAC		23
Pou5f1 (m)	F	GAGACTTTGCAGCCTGAGGG	60°C	20
	R	CTTTCATGTCCTGGGACTCCTC		22
Sox2 (m)	F	CAGGAGTTGTCAAGGCAGAGA	60°C	21
	R	CTTAAGCCTCGGGCTCCAAA		20
Klf4 (m)	F	AGAACAGCCACCCACACTTG	60°C	20
	R	GTGGTAAGGTTTCTCGCCTGT		21
Nanog (m)	F	AAAGGATGAAGTGCAAGCGG	59°C	20
	R	GTGCTGAGCCCTTCTGAATC		20
Vimentin (m)	F	AACGAGTACCGGAGACAGGT	60°C	20

	R	CAGGGACTCGTTAGTGCCTTT		21
Nestin (m)	F	ACATACAGGACTCTGCTGGAG	59°C	21
	R	CCAAGAGAAGCCTGGGAACT		20
Actin (m)	F	CGCCACCAGTTCGCCAT	60°C	17
	R	CTTTGCACATGCCGGAGC		18
RNMT (h)	F	TGAGTGTGACGGCTGGAACTC	60°C	21
	R	CACGCGTTGGGTAGTGTGAAG		19
RAM (h)	F	CCTCAAACCTTTGGGATT	50°C	18
	R	TTCCTGATACTCCTTGTC		18
POU5F1 (h)	F	CCCACACTGCAGCAGATCA	60°C	19
	R	ACCACACTCGGACCACATCC		20
SOX2 (h)	F	GCGGAAAACCAAGACGCTCAT	61°C	21
	R	CATGCTATTGCCGCCGGG		18
NANOG (h)	F	ACAGGTGAAGACCTGGTTCC	59°C	20
	R	GAGGCCTTCTGCGTCACA		18
PAX6 (h)	F	AGTGCCCGTCCATCTTTGC	60°C	19
	R	CGCTTGGTATGTTATCGTTGGT		22
GAPDH (h)	F	CTTCGCTCTCTGCTCCTCCT	61°C	20
	R	CGATGTGGCTCGGCTGG		17

2.4.3 Polysome profiling

To perform polysome profiling 2.5x 10^6 mESCs were seeded into a 10 cm dish and treated with non-targeting control and siRNA against RAM for 48 hr. Cells were replenished with fresh culture media 1 hr prior to cell lysis. Cells were incubated 10 min at 37°C with 100 µg/mL of cycloheximide prior to the start of polysome purification procedure. From now on all the steps were carried out on ice. Cells were quickly washed with ice-cold PBS containing 100µg/mL of

cycloheximide and collected by scraping with in Polysome Lysis buffer (Table 2.1).

The lysates were loaded on top of 10 mL sucrose gradients (2 mL of each gradient 10-50% were disposed on tubes in an increment order) and centrifuged at 38,000 rpm for 2 hr at 4°C in Beckman ultracentrifuge using a SW41 rotor. Gradients were fractionated (0.5 mL fractions) and each fraction was collected with a Foxy Jr. (Teledyne) fraction collector. RNA content from each fraction was purified using RNeasy GeneJet RNA purification kit (Thermo Scientific), quantitated and checked for purity by Nanodrop 1000 and used for RT-qPCR analysis. Ct values were normalised to input (RNA prior fractionation).

2.4.4 *In vitro* cap methyltransferase activity

For the *in vitro* cap methylation reaction, protein extracts (0.25 µg, 0.5 µg or 1 µg) lysed with F-buffer were combined with 200 µM S-adenosylmethionine and 1/50 purified capped transcripts in MT assay buffer (Table 2.1) in a final volume of 10 µL and incubated at 37ºC for 10 min unless indicated differently. Following the reaction, RNA was purified by phenol/chloroform extraction.

The pelleted RNA was resuspended in 4 µL of 50 mM sodium acetate (pH 5.5) and incubated with 1 U P1nuclease for 60 min at 37ºC to release free GpppG or m7GpppG. [α-32^P]–labelled GpppG, or [α-32^P]–labelled m7GpppG spots were resolved by thin layer chromatography (TLC) using PEI (polyethylenimine) cellulose plates in 0.4 M ammonium sulphate. The TLC was visualised using autoradiography and quantified using AIDA imager analyser.

2.4.5 Immunoprecipitation of methyl capped mRNAs

Total RNA from mESCs transfected with non-targeting control or with siRNA against RAM was extracted as described previously. The immunoprecipitation protocol was performed by Dr. Victoria Cowling. mRNA was purified from 100 µg of total RNA using standard laboratory protocol. 10 µL purified anti-m7G antibody or control antibody was pre-bound to 25 µL Protein A/G agarose in PBS, 0.01% Triton, 0.1 mg/mL BSA, 0.1 mg/mL polyU and 1 mM DTT for 30 min at room temperature followed by two washes. Immunoprecipitations were carried out using 2 µg oligo dT-purified RNA in 200 µL of the same buffer complemented with 5 U RNasin for 1 hr at room temperature followed by three washes. RNA was purified from the IPs by phenol/chloroform extraction and precipitated using 1/10th volume of 5 M sodium acetate and 4 µg tRNA as a carrier. RNA was resuspended in 50 µL water. RT-qPCR was performed to detect the methyl cap transcript levels of mRNAs of interest and Ct values were normalised to control methyl IP.

3 RNMT and RAM are highly expressed in embryonic stem cells

3.1 Introduction

The methylation of the mRNA cap represents a crucial checkpoint during eukaryotic gene expression through which the fate of certain transcripts can be regulated. In fact, it is well established that the synthesis of N7-methylguanosine cap at the 5’ end of mRNA affects several processes involved in RNA metabolism such as splicing, nuclear export, mRNA decay and silencing (Cheng et al., 2006; Topisirovic et al., 2011). In mammals, addition of the methyl cap is mediated by the sequential activity of two enzymes: the bi-functional guanylyltransferase/triphosphatase or RNGTT and RNA methyltransferase RNMT and its coactivator subunit RAM (Furuichi & Miura, 1975; Gonatopoulos-Pournatzis et al., 2011). The formation of the RNA methyl cap is so critical for viability and the relative functions of mRNA that it has long been considered to be a constitutive process within cells. Only recently, accumulating evidence has suggested that this modification is subject to regulatory mechanisms. Indeed pioneering studies have not only revealed that environmental conditions can affect the cellular levels of cap methylation but also that this regulation may be specific for certain transcripts (Chang et al., 2012; Fernandez-Sanchez et al., 2009)

One of the most fascinating scenarios where the gene expression machinery needs to be regulated to quickly adapt to the cues that cells receive from the environment is during stem cell fate determination. An open chromatin structure, the presence of “bivalent” domains and the activity of super enhancers are some of the characteristic features of the unique gene expression landscape within stem cells (Bernstein et al., 2006; Efroni et al., 2008; Meshorer et al., 2006; Whyte et al., 2013).

Moreover, although controversial data have been reported, recent evidence is directly correlating the most abundant internal modification of mRNA, the N6-methyladenosine (m^6A) with cell pluripotency. Despite having opposite results, both studies showed that alterations in the expression levels of the methyltransferase responsible for the formation of m^6A, cause perturbations in the ability of cells to self-renewal either increasing or diminishing it (Batista et al., 2014; Wang et al., 2014b). Furthermore, it was recently reported that in ESCs, the THO nuclear export complex preferentially recognises and binds mRNAs expressing pluripotency factors mediating their nuclear export and consequently their efficient translation. Depletion of components of the THO complex led to the loss of self-renewal (Wang et al., 2013). Taken together these findings strengthen the role that post-transcriptional mechanism such as RNA modifications and RNA binding proteins play in the maintenance of pluripotency (Guallar & Wang, 2014; Kwon et al., 2013).

Considering the aforementioned data, the study of mRNA cap methylation in stem cells may reveal novel insight into the specific gene expression program required for the maintenance of pluripotency. This chapter will be focused on investigating the expression and activity of RNA methyltransferase (RNMT) and its interacting partner RAM in embryonic stem cells.

3.2 Results

3.2.1 Characterization of the mESC line 46C.

To accomplish the aims of this chapter, a mouse embryonic stem cell line cultured in media supplemented with LIF was established for the first time in the

Cowling lab. The Sox1:GFP knock-in (46C) ESCs were employed in this work and were kindly donated by Dr. Marios Stavridis (University of Dundee), who regularly provides advice and suggestions. It is extremely important that stem cells are constantly and unequivocally maintained in an undifferentiated state during the whole experimental process. To verify the cell pluripotency, three independent assays were performed periodically and in concomitance of every experiment. Alkaline phosphatase (AP) assay was used as a first method to assess cell pluripotency. It is well described that the high activity of AP on the surface of undifferentiated cells makes the AP a suitable marker of pluripotent stem cells (Marti *et al.*, 2013). Briefly, mESCs were plated on a 6-well dish and after three days cells were fixed and stained to reveal alkaline phosphatase activity. The activity levels of AP were compared in mESCs and in immortalised mouse embryonic fibroblasts (MEFs), used as a negative control, which were kindly donated by Ian Ganley's lab. The presence of purple colonies confirmed the pluripotency of the mESCs and the concomitant absence of staining in MEFs ensured the specificity of the assay (Figure 3.1a).

Extensive studies on stem cells have established the main transcription factors required for pluripotency (Avilion et al., 2003; Niwa et al., 2000; Chambers et al., 2003). Thus, the second assay used to evaluate pluripotency involved the analysis of protein and transcript levels of these previously defined pluripotency factors. Sox2 and Oct4 protein expression was found to be abundant in mESCs but was undetectable in MEFs (Figure 3.1b). Further validation came from RT-qPCR analysis, which revealed that mRNA levels of Sox2, Oct4, Nanog and Klf4 were tenfold higher in stem cells compared to MEFs (Figure 3.1c).

Vimentin, which is the most frequently found intermediate filament protein in fibroblasts (Franke *et al.*, 1982), was used as positive control for MEFs.

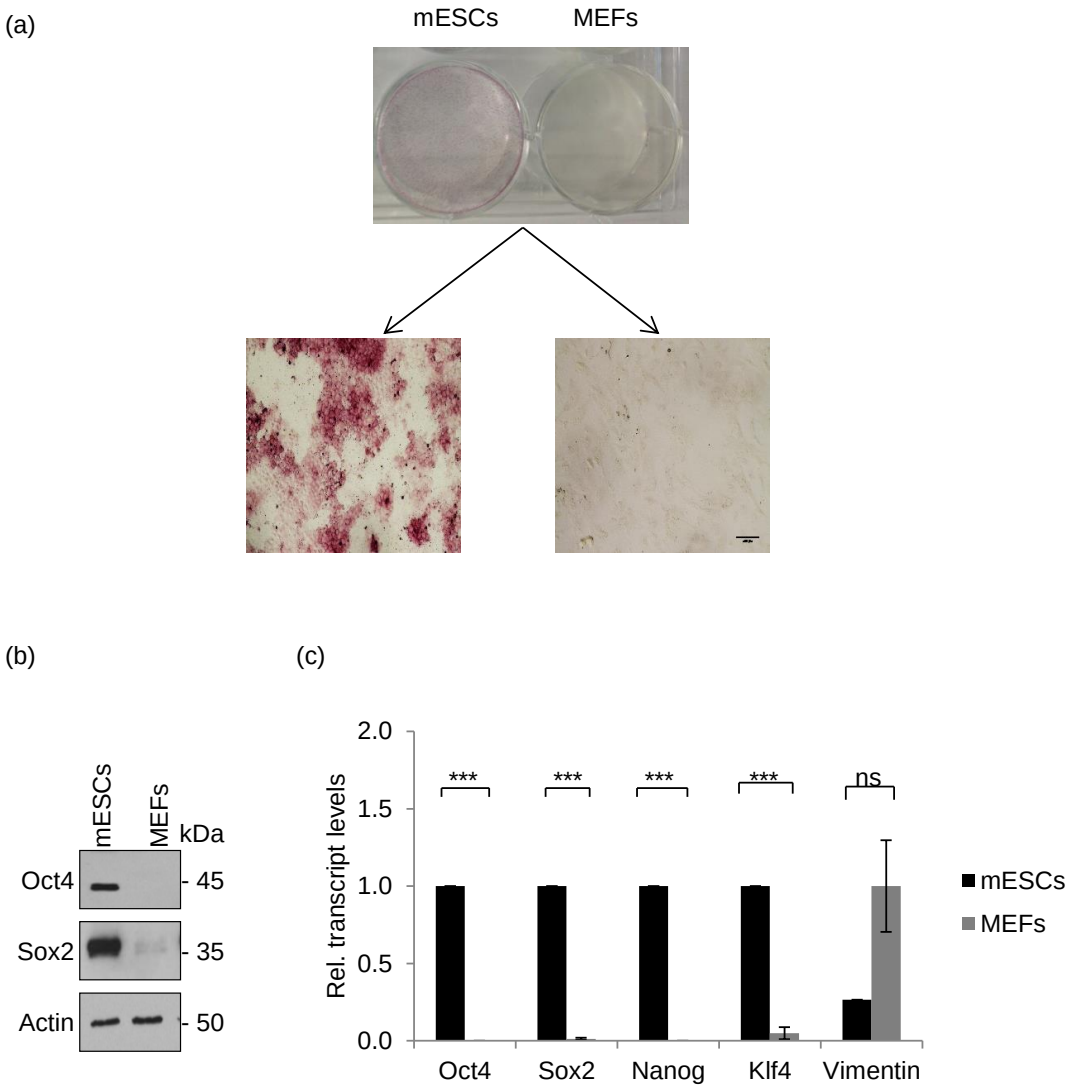


Figure 3.1: Three different assays to characterise the pluripotency of the mouse embryonic stem cells.
(a) Alkaline phosphatase expression in mouse embryonic stem cells (mESCs) and immortalised mouse embryonic fibroblast (MEFs). Cells were seeded into a 6 cm dish and after three days were stained to reveal alkaline phosphatase activity. Images were taken with Canon camera EOS 100D. Scale bar measures 500 µm. (b) Cell extracts from mESCs and immortalised MEFs were analysed by Western blot with antibodies against the indicated proteins. Actin serves as the loading control. (c) RNA extracted from mESCs and MEFs was analysed by RT-qPCR. The data were normalised to Actin mRNA levels. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). * indicates $p< 0.05$, ** indicates $p<0.005$, *** indicates $p<0.001$ and ns indicates $p>0.05$. Bar charts depicts the average values and relative standard deviation of three independent experiments. All the above assays are representative of three biological replicates.

Once the stem cells were characterized, the next step was to test the specificity of anti-RNMT and anti-RAM antibodies. Previously, Cowling laboratory has only used human cell lines and consequently antibodies were only raised against the human orthologues of RNMT and RAM. In order to work with mESCs, antibodies that recognise the mouse homologue were required. The existing anti-RAM antibodies were found to recognise a band of the appropriate size in Western blot analysis, whereas the existing RNMT antibody did not. This could be because human and mouse RNMT homologues differ in the N-terminal domain. Thus an antibody against mouse RNMT was raised by the University of Dundee DSTT (Division of Signal Transduction Therapy) service. Briefly, recombinant GST-RNMT was injected into a sheep and 20 weeks later blood serum was collected and anti-RNMT antibodies were affinity purified. The specificity of the generated anti-RNMT, and that of the already available anti-RAM antibody, was determined by depleting the enzymes from mESCs using RNAi. Cells were treated for 48 hr with three independent siRNAs against RNMT, three independent siRNAs against RAM or with a non-targeting siRNA (Ct) used as negative control. Clearly the depletion of RNMT and RAM caused the loss of the respective bands observed by Western blot analysis confirming the specificity of the antibodies tested (Figure 3.2). This also confirmed that the siRNAs successfully reduced protein expression and thus can be further employed in this study.

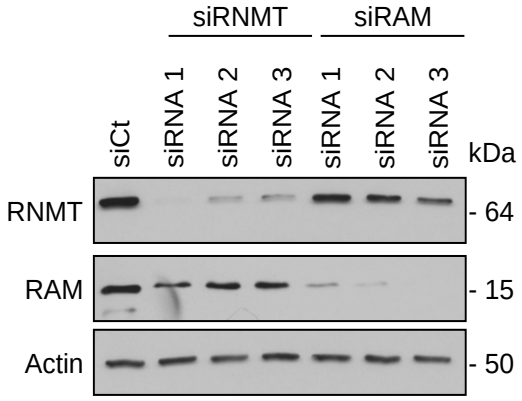


Figure 3.2: Characterization of anti-RNMT and anti-RAM antibody specificities in mESCs using three different siRNAs.

mESCs were transfected with three independent siRNAs against RNMT, three independent siRNA against RAM (1, 2, and 3) and a non-targeting control (siCt) for 48 hr. Cell extracts were analysed by Western blot with antibodies against the indicated proteins. Actin serves as loading control. The above assays are representative of two biological replicates.

3.2.2 RNMT and RAM expression in mESCs

In order to address the aims of the project, the expression levels of the enzymes involved in the methylation of the mRNA cap were investigated for the first time in mESCs. The expression of RNMT and RAM in mESCs was compared to immortalised MEFs, a terminally differentiated cell line (Figure 3.3). Both RNMT and RAM protein levels were found to be elevated in mESCs compared to MEFs, which was also mirrored by the mRNA levels. In fact, the mRNA levels of RNMT and RAM also revealed that the two proteins are also transcriptionally regulated. Concerning the transcript data, RNMT primers were designed against the most abundant splicing variants whereas RAM only presents one splicing variant. The efficiency of the primers was also tested and they were found to work specifically.

This results documented the abundance of RNMT and RAM expression in mESCs and raised the question whether the overall methyltransferase activity could be higher in mESCs than in MEFs. To address this, an *in vitro* methyltransferase activity was performed, where the cap methyltransferase activity within titration of mESC and MEF whole cell extracts was measured in a 5 min reaction with ^{32}P-capped transcript and S-adenosylmethionine (SAM). Following the reaction, the capped transcript was nuclease digested and the resultant GpppG (cap) and m^{7}GpppG (methyl cap) were resolved by thin layer chromatography (Figure 3.4). The amount of m^{7}GpppG are plotted as percentage of total cap (i.e. m^{7}GpppG x 100/ (m^{7}GpppG + GpppG)) relative to amount of protein lysate used. As expected, higher cap methyltransferase activity was observed in mESCs compared to MEFs. In particular, the maximal difference in methyltransferase activity between mESCs and MEFs was

threefold. Overall these results show that not only the expression of methyltransferase complex is higher in mESCs, but also its activity is substantially elevated when compared to MEFs.

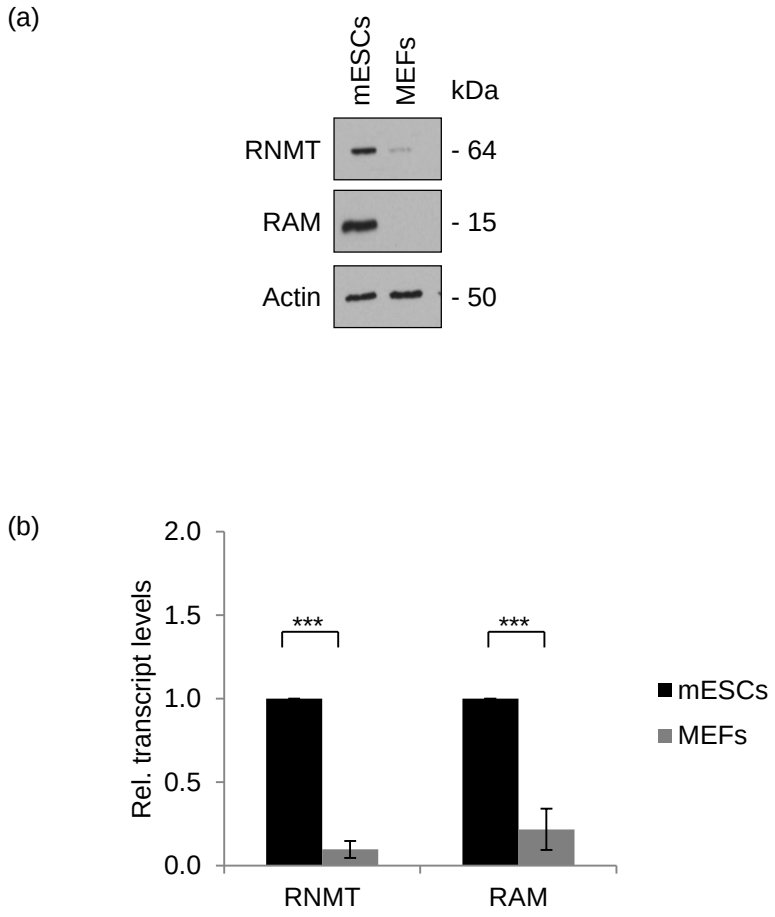


Figure 3.3: RNMT and RAM show elevated expression in mESCs compared to immortalised MEFs.
In vitro cap methyltransferase activity assay. (a) Cell extracts from mESCs and in immortalised MEFs were analysed by Western blot with antibodies against the indicated proteins. Actin serves as the loading control. (b) RNA extracted from mESCs and MEFs was analysed by RT-qPCR. The data were normalised to Actin mRNA levels. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). *** indicates $p<0.001$. Bar charts depicts the average values and relative standard deviation of three independent experiments. All the above assays are representative of at least three biological replicates**.**

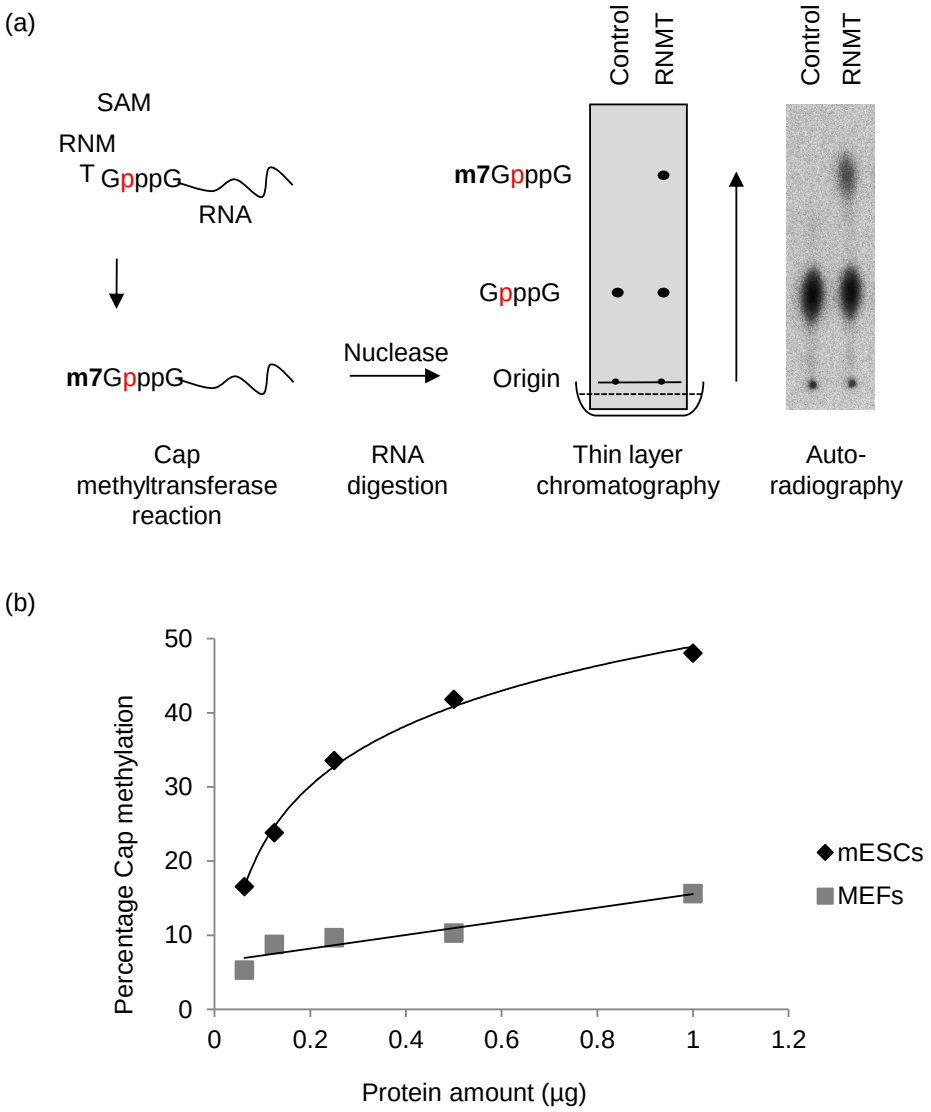


Figure 3.4 *In vitro* cap methyltransferase activity assay.
(a) RNMT and RAM activity was assessed in an *in vitro* cap methyltransferase activity assay. This schematic diagram represents the major steps of the assay. A typical picture of an autoradiography exposure is shown on the right. (b) Relative cap methyltransferase activity was detected in cell extracts using an *in vitro* cap methyltransferase assay. Activity was measured for 5 min with a titration of cell extracts. Quantification of thin layer chromatography was performed using ImageJ densitometry and the activity expressed as percentage methylation (i.e. (m^7GpppG x 100) / (m^7GpppG + GpppG)). The above assay is representative of two biological replicates.

To discard the hypothesis that the immortalisation of MEFs could have somehow affected the expression of RNMT and RAM, mice of 129/Ola strain (same mice from which mESCs used in this work were derived) were used to isolate primary MEFs. The experiments were repeated with primary MEFs and confirmed that RNMT and RAM expression in mESCs is indeed higher than in MEFs, both at protein and transcript levels (Figure 3.5). In all the replicates, expression of RAM in primary MEFs was slightly higher compared to RAM levels observed in immortalised MEFs as well as Actin levels. Therefore, combing these observations together it emerged that the immortalisation process has indeed affected RAM expression. Nevertheless, the results between immortalised and primary MEFs consistently showed that the expression both at protein and transcripts levels of RNMT and RAM is higher in mESCs than MEFs. The concomitant analysis of pluripotency markers validated the pluripotent status of stem cells employed in the study.

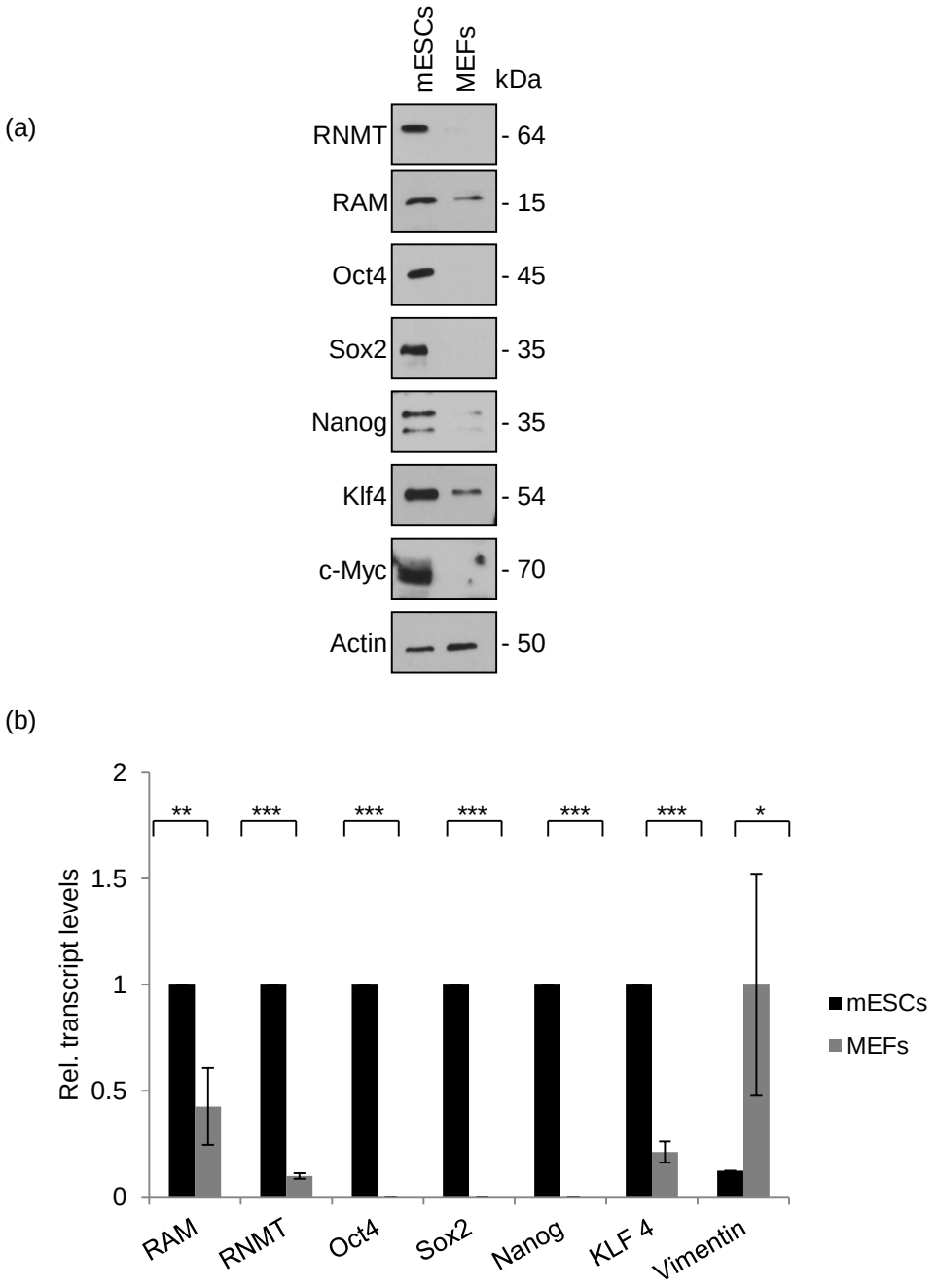


Figure 3.5: RNMT and RAM are highly expressed in mESCs in comparison to primary MEFs.

(a) Cell extracts from mESCs and primary MEFs from 129 Ola mice were analysed by Western blots with antibodies against the indicated proteins. Actin serves as the loading control. (b) RNA extracted from mESCs and primary MEFs was analysed by RT-qPCR. The data were normalised to Actin mRNA levels. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). * indicates $p< 0.05$, ** indicates $p<0.005$ and *** indicates $p<0.001$. Bar charts depicts the average values and relative standard deviation of three independent experiments. All the above assays are representative of at least three biological replicates.

The aforementioned data shed the light on a differential expression of enzymes involved in methyl cap mRNA in mESCs. These observations encouraged further characterisation of the RNMT and RAM complex in mESCs. With this purpose, gel filtration or size exclusion analysis was performed to analyse the elution pattern of the RNMT and RAM complex (Figure 3.6a). 1 mg of extracts from mESCs were separated through a Superose 6 column, 24 fractions were collected and the elution profile was analysed by Western blot (Figure 3.6b). As previously reported in our laboratory, co-elution of RNMT and RAM was detected in fractions 6-7. The finding that RNMT and RAM complexes migrate approximately at 150 kDa, which is higher than the expected size, suggests that may be part of a higher order complex and this may also explain the presence of the two proteins immediately after the void volume (fractions 1 and 2). However, differently from previous data, RNMT was individually detected in fraction 8 and 9. This could indicate the presence of monomeric RNMT in mESCs. To collect more evidence, anti-RNMT or anti-RAM antibodies were used to immunodeplete the relative proteins from 1 mg of mESCs extract. The experimental conditions were sufficient to totally and successfully immunodeplete RAM following RNMT immunoprecipitation (IP) (Figure 3.6c). On the contrary, RNMT was clearly detectable in the flow through following RAM IP indicating that not all RNMT may interact with RAM in mESCs. Hence, the gel filtration and the immunodepletion experiments suggest that RNMT in mESCs may be present independently from RAM. The previous findings support further investigation of RNMT and RAM expression during development to see whether the expression of this complex changes following differentiation.

For this purpose, protein extracts from different mouse organs were analysed by Western blot (Figure 3.7). Organs from two mice were processed and extracts kindly provided by Dr. Francisco Inesta Vaquera. The results between the two mice are consistent and showed firstly, that the individual expression of RNMT and RAM changes dramatically amongst different mouse organs. Secondly, RNMT and RAM stoichiometry is variable between the different organs. For instance, RNMT levels in some organs were higher than RAM (e.g. the brain) and *vice versa* (e.g. the heart). These data together indicate that the expression levels of the cap methylation machinery can vary drastically following differentiation depending on the cell lineage specification.

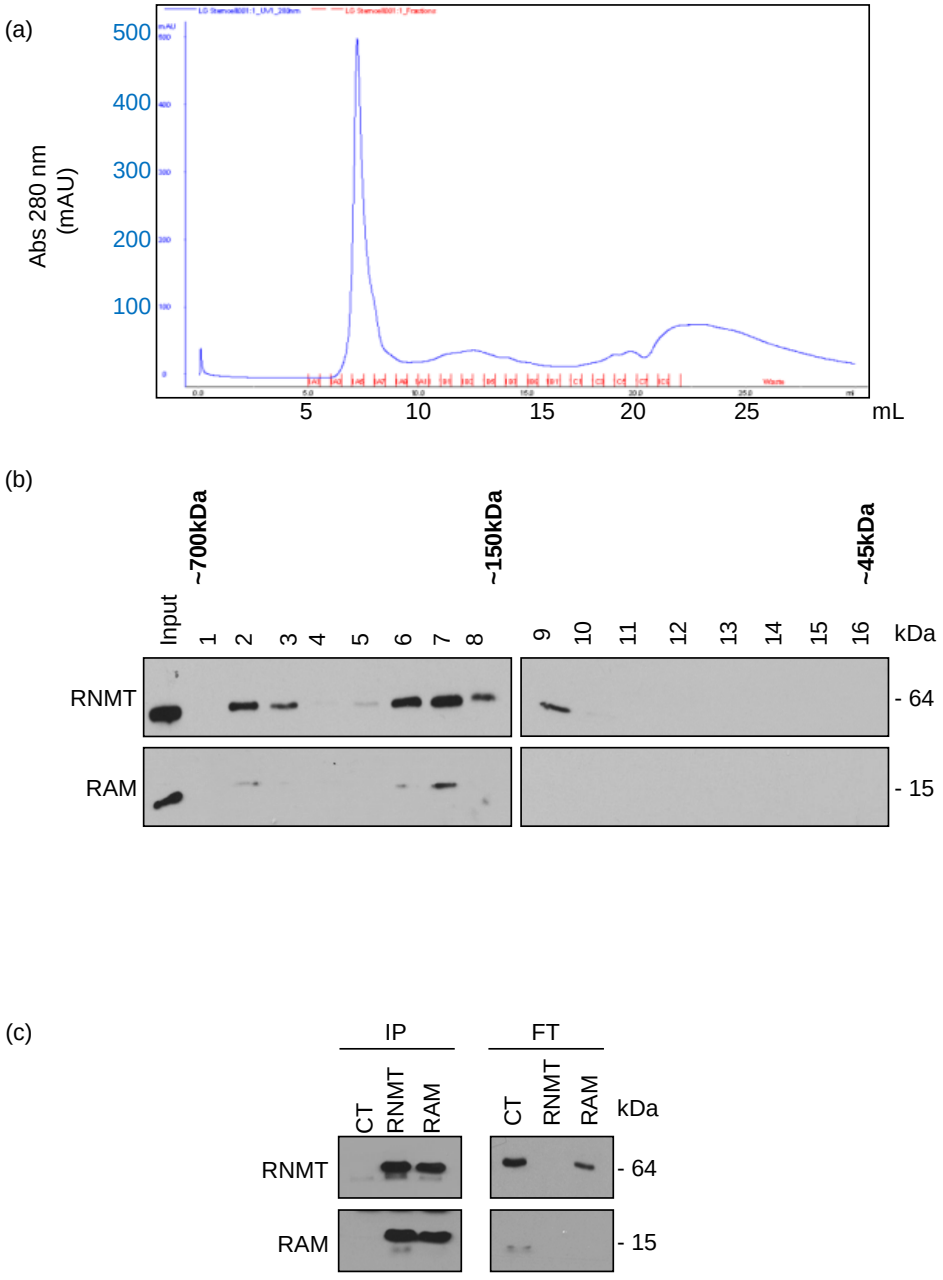


Figure 3.6: Not all RNMT is bound to RAM in mESCs.
(a) Elution profile of size exclusion chromatography. The blue line represents the UV absorbance at 280 nm. (b) 1 mg of mESCs cell extracts were resolved through a Superose 6 column and fractions collected and analysed by Western blot with antibodies against the indicated proteins. (c) 1 mg of mESCs cell extracts were incubated with 1 µg of anti RNMT or anti RAM antibody for 2 hr at 4°C analysed by Western blot with antibodies against the indicated proteins. All the above assays are representative of two biological replicates.

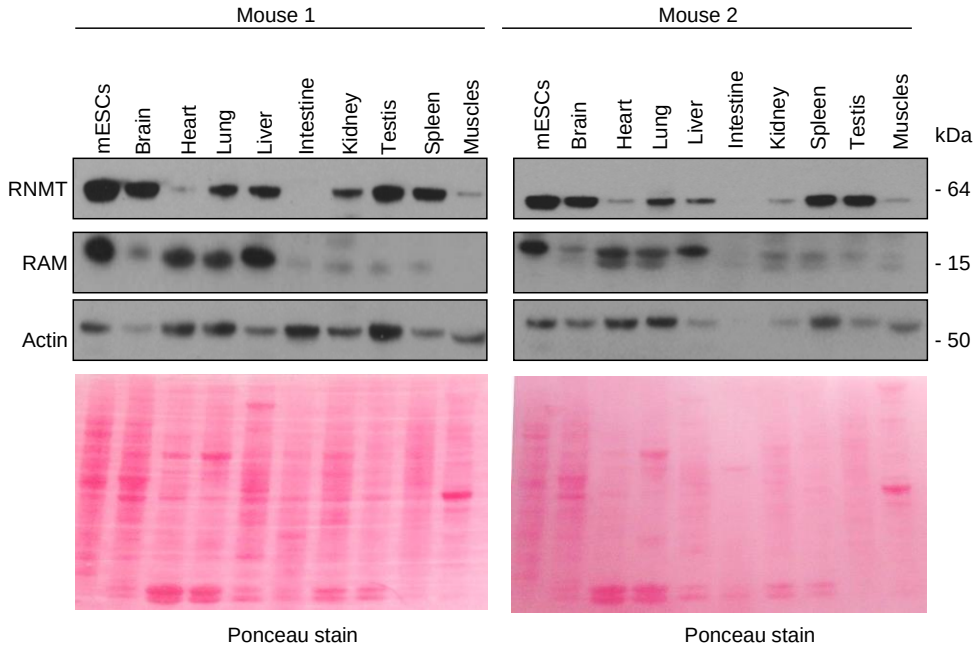


Figure 3.7: RNMT and RAM are differentially expressed amongst different mouse organs.

Protein extracts (30 µg) from different mouse organs listed above were analysed by Western blot with antibodies against the indicated proteins. The analysis was performed with 5 weeks old mice. Ponceau staining as the loading control. Tissue samples were prepared by Dr. Francisco Inesta-Vaquera. The above assays are representative of two biological replicates.

3.2.3 RNMT and RAM in reprogramming

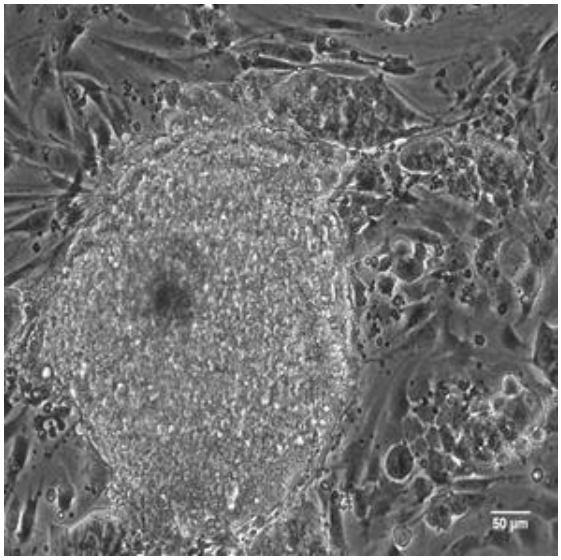
In the last decade the major breakthrough in stem cells biology was the reprogramming of a differentiated cell to an induced pluripotent stem cell (iPSC) (Takahashi & Yamanaka, 2006). Nowadays the generation of iPSCs is a common protocol used to investigate whether the protein under investigation is involved in the establishment of the pluripotent state. In figure 3.3 and 3.5, it was observed that RNMT and RAM expression was lower in MEFs (both immortalised and primary) than in mESCs. With the intention to establish whether the expression of the two proteins was restored when MEFs were reprogrammed, different protocols were tried to make iPSCs from MEFs.

For the first protocol, transduction of MEFs with lentivirus expressing Oct4, Klf4, Sox2 and c-Myc (OKSM) factors, kindly donated by Dr. Lindsay Davidson (University of Dundee), was attempted. This transduction method was first tried in a 96 well format with HEK293 cells and despite being successful, the process of scaling up the protocol on MEFs was not (data not shown). Subsequently, Dr. Marios Stavridis suggested and kindly donated reagents for a protocol in which MEFs were transfected with plasmid containing the defined transcription factors (Oct4, Sox2, Klf4, and c-Myc) (OSKM), fused in frame via 2A sequences and co-expressed as a single origin of replication (ORF). Optimization of the protocol included seeding different cell densities and trying different media compositions.

Briefly, MEFs were plated in a well of a 6 well plate and were transfected with the plasmid containing OSKM the day after. Selection of transfected cells was

performed with doxycycline on the following day. After 4 days of selection, alterations in cell morphology were observed whereby cell shape changed from the typical stretched fibroblast shape to a more rounded circular one. Group of cells that were successfully reprogrammed started to assume a colony morphology that became evident after 15 days (Figure 3.8a). These colonies were big enough to be picked and cultured to make independent iPS cell lines. Once the cell line was established, cells were seeded and stained for AP activity to verify the acquired pluripotency. The purple staining of iPSC colonies confirmed that the reprogrammed cells were pluripotent. For this assay, MEFs provided the negative control, whereas mESCs were the positive one (Figure 3.8b). The high expression levels of Oct4, Sox2, Nanog, c-Myc and E-cadherin further confirmed the pluripotent state of mESCs and mouse iPSCs. The band indicated to recognise c-Myc has been verified by c-Myc siRNA transfection. Vimentin expression identified MEFs as fibroblast (Figure 3.9a). Remarkably, the expression of RNMT and RAM was totally restored in mouse iPSCs to similar levels observed in mESCs. The upregulation of RNMT and RAM following reprogramming was also confirmed at mRNA levels by RT-qPCR (Figure 3.9b) whereby mESCs and mouse iPSCs, apart from exhibiting high levels of pluripotency factors, were characterized by high RNMT and RAM transcript levels. Also in this case, Vimentin was exclusively highly expressed in MEFs.

(a)



(b)

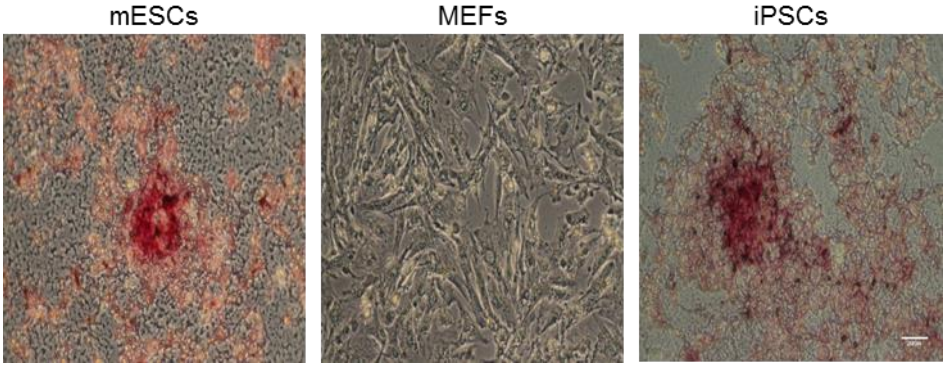


Figure 3.8: Derivation of induced pluripotent stem cells.
Primary MEFs were reprogrammed by transfection with plasmid containing defined transcription factors (TFs) (Oct4, Sox2, Klf4, and c-Myc) fused in frame via 2A sequences and co-expressed as a single ORF. (a) Bright field image of an iPS colony at 20x magnification. Scale bar measures 50 µm. (b) Bright field image of mESCs, MEFs and iPS stained to reveal alkaline phosphatase activity. Scale bar measures 200 µm.

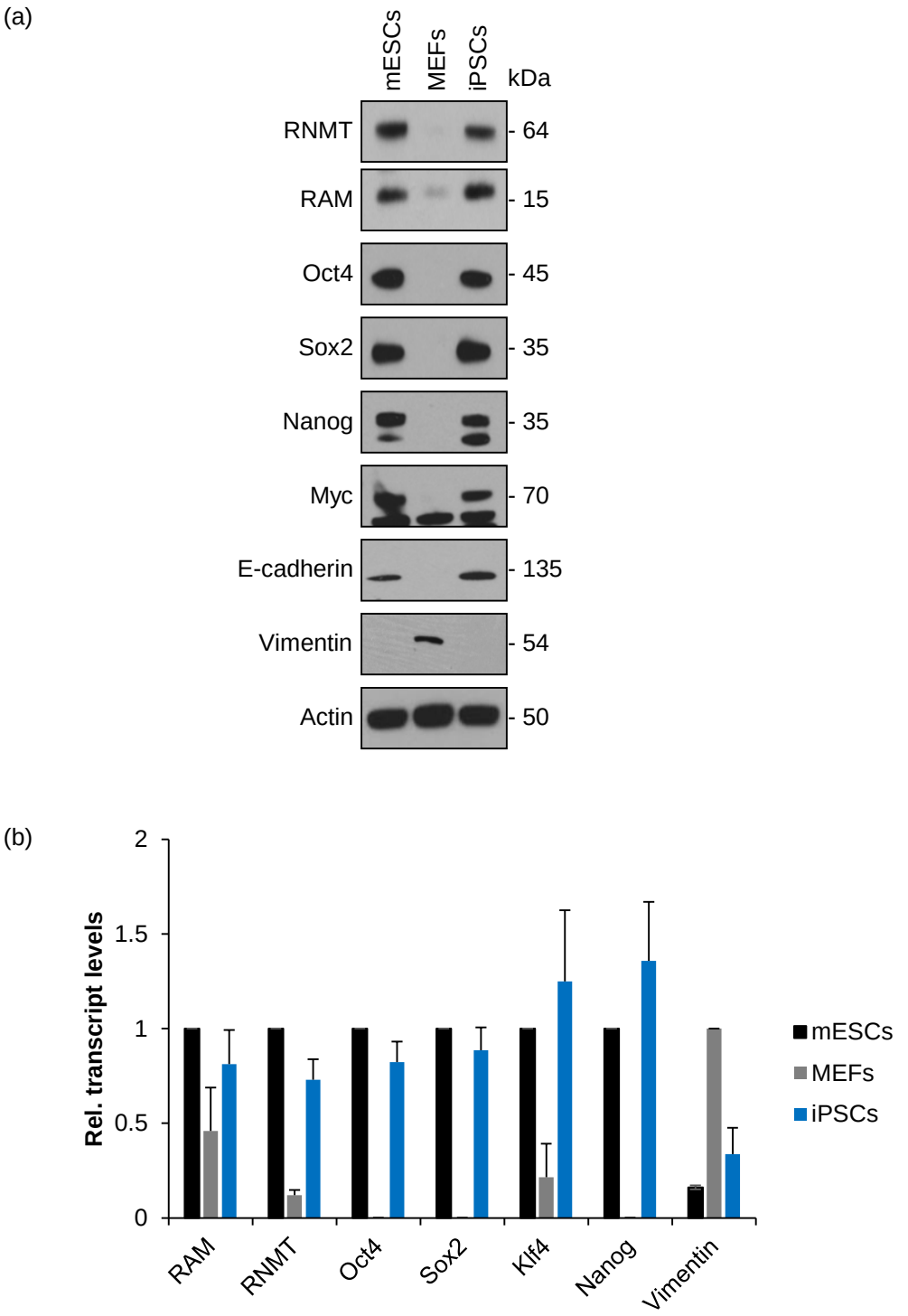
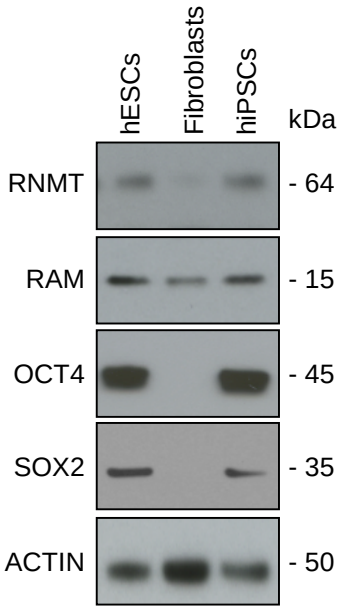


Figure 3.9: RNMT and RAM expression levels are restored in iPSCs.
Cell extracts from mESCs, MEFs and mouse iPSCs were analysed by Western blot with antibodies against the indicated proteins. Actin serves as the loading control. (b) RNA extracted from mESCs, MEFs and mouse iPSCs was analysed by RT-qPCR. The data were normalised to Actin mRNA levels. The above analyses are representative of three biological replicates.

Even more importantly, when the same analysis was carried out with human ESCs, human fibroblast and human iPSCs, the same pattern of RNMT and RAM expression was observed as with the mouse cell lines (Figure 3.10). In fact, the pluripotent hESCs and hiPSCs, characterized by the elevated levels of pluripotency factors, presented higher expression at both transcript and protein levels of RNMT and RAM compared to fibroblasts. The three human cell lines were kindly provided by Dr. Lindsay Davidson. The data presented so far clearly show that high levels of RNMT and RAM are a feature of pluripotent cells, whereas their expression is reduced in terminally differentiated MEFs.

(a)



(b)

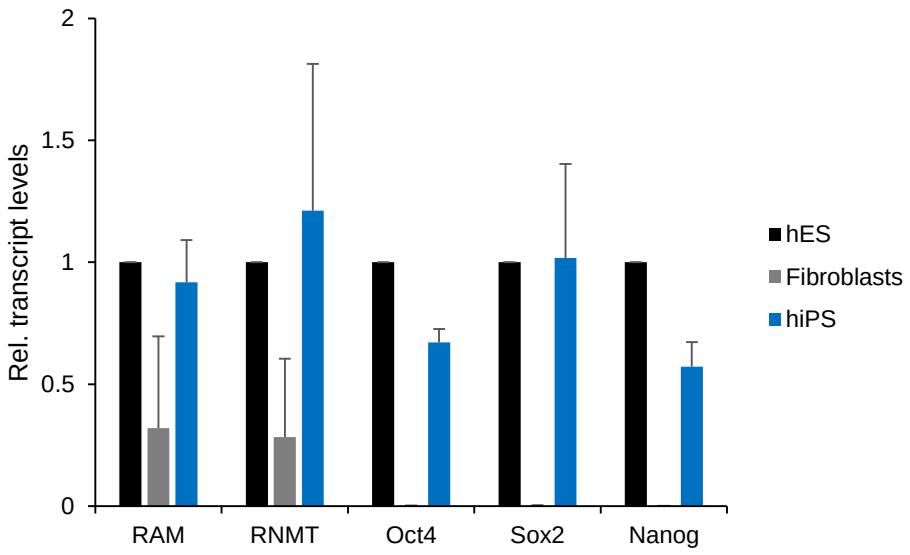


Figure 3.10: RNMT and RAM expression levels are restored in human iPSCs.
(a) Cell extracts from human ESC 181 cell line (hESCs), human fibroblasts and human induced pluripotent stem cells (hiPSCs), kindly donated by Dr. Lindsay Davidson, were analysed by Western blot with the antibodies against the indicated proteins. Actin serves as the loading control. (b) RNA extracted from human ESCs, human fibroblast and human iPSCs was analysed by RT-qPCR. The data were normalised to Actin mRNA levels. Samples were kindly provided by Dr. Lindsay Davidson. The above analyses are representative of three biological replicates.

3.3 Discussion

Considering that addition of the cap, its methylation and its eventual removal are fundamental processes that control RNA metabolism in mammalian cells, we reasoned that the regulation of cap methylation might be critical for the maintenance of pluripotency. With the aim to understand whether mRNA cap methylation represents a novel regulatory process in stem cell transcription, a stem cell line was established for the first time in the Cowling laboratory. Characterisation of the cell line confirmed cell pluripotency and thus the reliability of the cell system. The speculation that the enzyme responsible for cap methylation could play a role in maintenance of pluripotency was strengthened by data comparing the expression of the methyltransferase RNMT and its co-factor RAM in mESCs and MEFs. Although the Western blot is not a quantitative analysis, the dramatic difference in RNMT and RAM expression, compared to primary and immortalised MEFs, strongly supported the conclusion that the two proteins are highly expressed in mESCs. As expected, despite exhibiting lower methyltransferase activity compared to mESCs, MEFs were able to convert the guanosine cap to methyl guanosine cap. This observation reinforced the assumption that the methylation of the cap, being so critical for mRNA half-life, is a fundamental process and despite the proteins responsible for the modification could be less abundant, they are ubiquitously present in cells and consequently active. This result led to assume that the higher enzymatic activity may be specifically required to face the high global transcription levels that characterize stem cells (Efroni *et al*., 2008) whereas the basal RNA methyltransferase activity is sufficiently catalysed by lower levels of RNMT and RAM.

The high level of RNMT and RAM expression in mESCs encouraged further investigation into the nature of the RNMT and RAM complex by gel filtration and IP assays. RNMT, but nor RAM, was eluted in fractions 8 and 9 of gel filtration analysis and the data was corroborated by the immunodepletion experiment whereby RNMT signal was found in the flow through following RAM IP. Therefore, both assays pointed out that not all RNMT might be bound to RAM in mESCs. This is surprising because in all the cell lines previously investigated, RNMT has always been seen in a complex with RAM and monomeric RNMT or RAM have never been detected. Although previous findings on the equimolar RNMT and RAM complex have led to the assumption that the two proteins are co-regulated, the data described above prompted us to investigate whether the two proteins are differently expressed in the mouse organs. Western blot analysis showed that RNMT and RAM expression is organ-specific as their levels are differently regulated in different organs. Moreover, the variability in the detection of RNMT and RAM in different organs supports the hypothesis that RNMT and RAM may exist as monomers or in different complexes in different cell types. Consistently to this, the individual depletion of RNMT or RAM in mESCs did not significantly alter the expression of the other (Figure 3.2), which is contrary to what has been previously observed in other cell lines whereby following RNMT depletion by RNAi, RAM expression was also impaired and *vice versa* (Gonatopoulos-Pournatzis et al., 2011).

These finding together show that RNMT and RAM expression in mESCs may be differently regulated than previously described and raise the question whether these two proteins fulfil different roles within the cap methylation

process. The reported findings are relevant in light of the aforementioned work which identified the THO complex, the nuclear protein complex, as a novel regulator of pluripotency (Wang et al., 2013). Wang and colleagues hypothesised that the recognition mechanism is due to the presence of the methyl cap that could specifically be present on mRNA of pluripotency factors. These data could potentially explain that high level of RNMT and RAM in mESCs preferentially methylates certain transcripts to make them available for the further steps of gene expression.

Lastly, to assess whether high levels of RNMT and RAM complex are a feature of all pluripotent cells, MEFs were reprogrammed to iPSCs and expression of the complex was investigated. The reprogramming caused the expression of RNMT and RAM to be restored to levels observed in mESCs. Excitingly, the same results were observed in human ESCs and human iPSCs. This clearly shows that the regulation pathway of RNMT and RAM in stem cells is conserved from mice to humans. In order to discern whether RNMT and RAM are required for pluripotency or their expression is consequently upregulated following acquisition of pluripotency, in the future the transcript and protein levels will be analysed during the reprogramming process. Furthermore, to know how important RNMT and RAM expression are for reprogramming, the efficiency of the process will be assessed counting the iPS colonies following overexpression or depletion of RNMT and RAM. Data described above strongly encouraged us to carry on with the study of methyl cap mRNA in stem cells because any findings in mESCs could possibly be validated in hESCs and therefore having significant therapeutic implications.

4 RAM is regulated during neural differentiation

4.1 Introduction

One of the multiple reasons to study the stem cell biology is the possibility to closely investigate the phenomena that occurs during early embryo development. Additionally, ESCs possess the ability to generate a wide range of differentiated cell types (Smith, 2001) and this feature prompted investigation of different protocols to terminally derive pluripotent stem cells into specific cell lineage.

To date, one of the best characterised protocol to establish the conversion of mESCs into a differentiated cell type is represented by the *in vitro* neural differentiation. Previous studies identified Sox1 as an early neuroectodermal marker highly expressed in dividing neuroepithelial cells. Both *in vivo* (Sox1 is expressed in the neural tube) and *in vitro*, the expression of this transcription factor is related with the acquisition of neural fate and it is gradually lost during neuronal and glial differentiation (Pevny *et al.*, 1998). Based on these evidence, Ying and colleagues have engineered the 46C ES cells to express Sox1:GFP, where the open reading frame of the *Sox1* gene was replaced by GFPiresPac, and developed a protocol to differentiate the cells into neuroepithelial precursors. Authors were able to derive neural precursors by culturing the reporter cell line on gelatin coated plates, in serum free media and following LIF withdrawal. The Sox1:GFP^{+} neural precursors could further give rise to neurons and glial cells (Ying *et al.*, 2003b). Since then, the Sox1:GFP knock-in 46C ES cells have been widely used to investigate the kinetics of neural progenitor derivation.

Moreover, Abranches and colleagues were able to differentiate mESCs into neural progenitors in a rosette-like structure that recapitulates the neural tube formation. These evidence strongly support the similarity between the *in vitro* monolayer differentiation and the embryonic neural development (Abranches *et al.*, 2009).

In the previous chapter, it was reported that RNMT and RAM are expressed at higher levels in pluripotent stem cells than terminally differentiated cells. These data combined with the findings that RNMT and RAM levels vary among different organs, strongly encourage further studies to uncover the mechanism that drives RNMT and RAM expression during cell fate specification. Considering the brain tissue data and the simplicity of the monolayer protocol, we reasoned to closely study the regulation of RNMT and RAM expression levels during neural derivation. Therefore this chapter will be focused in the *in vitro* neural differentiation of mESCs.

4.2 Results

4.2.1 RNMT and RAM levels in neuronal cells

To begin with, RNMT and RAM expression was investigated in three different neuronal cell types. Primary cortical neurons and astrocytes were extracted, lysed and kindly donated by Dr. Ritchie Williamson (Daan van Aalten lab, University of Dundee). Astrocytes were established from post-natal day 2, while primary cortical neurons were isolated from 16-day old mouse embryos (Figure 4.1a). The expression of both RNMT and RAM was found to be dramatically reduced in astrocytes compared to mESCs, on the contrary, primary cortical

neurons retained similar levels of RNMT but low levels of RAM expression. To further confirm the abundance of RNMT and RAM in mESCs compared to differentiated cells, Dr Ritchie kindly provided other neuronal cells such as hippocampal neurons. Hippocampal cells were isolated from 16-day old mouse embryos and were cultured *in vitro* for the time indicated to obtain a fully polarized and mature phenotype (Williamson *et al.*, 2002) (Figure 4b). Fresh mESCs were run with astrocytes (sample was rerun) and hippocampal cells to directly compare the different cell types. The western blot analysis corroborated the data obtained so far whereby the protein levels of RNMT and RAM is clearly abundant in mESCs and poorly present in differentiated cells. In the fresh mESCs RNMT expression was higher than that in Figure 4.1a possibly due to longer incubation with chemiluminescent reagent whereas in astrocytes possibly the freezing-thawing process impoverished the already weak expression of RNMT, although a weaker band was still visible. The composition of lysis buffer used by Dr. Ritchie Williamson is very similar to the lysis buffer employed in this thesis. Therefore differences between samples were not amenable to different buffer composition. These results firstly confirm the higher levels of RNMT and RAM expression in mESCs compared to differentiated cells and secondly, it suggests that in the process of neural differentiation and consequential down regulation of RNMT and RAM proteins, RAM expression may be lost earlier compared to RNMT.

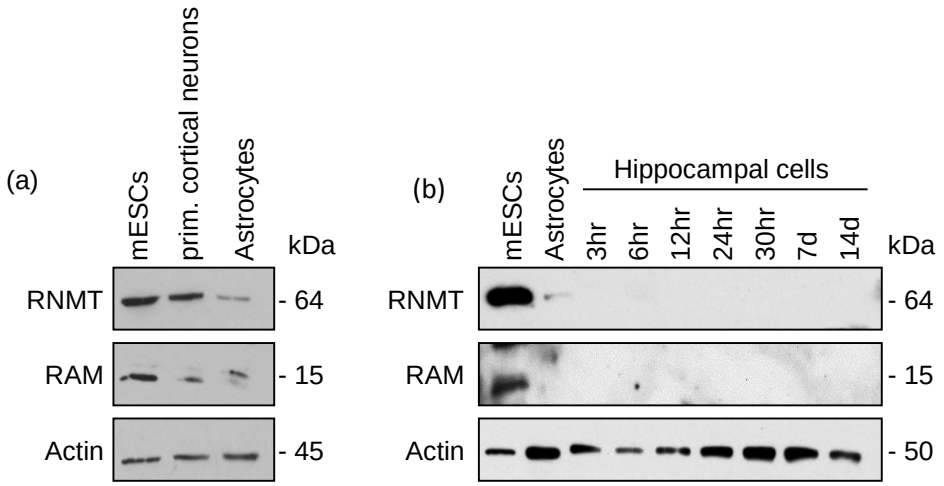


Figure 4.1: RNMT and RAM expression compared between mESCs and primary cortical neurons, astrocytes and hippocampal cells.
RNMT and RAM expression levels were analysed by Western blot performed on protein extracts from mESCs, primary cortical neurons (a), astrocytes (a) and hippocampal cells (b) with antibodies against the indicated proteins. Astrocytes were established from post-natal day 2. Primary cortical neurons and hippocampal cells were isolated from 16-days old mouse embryos, the latter were cultured *in vitro* for the time shown before being used for the analysis. Actin serves as the loading control. Primary cortical neurons, astrocytes and hippocampal cells were extracted, lysed and kindly donated by kindly donated by Dr. Richard Williamson (Daan van Aalten lab, University of Dundee). The above assay is representative of one biological replicate.

4.2.2 *In vitro* neural differentiation

In order to trace when exactly RNMT and RAM levels change during the differentiation process, mESCs were differentiated *in vitro*. As previously mentioned, the mESCs employed in this study were the Sox1:GFP knock-in (46C) ESCs, expression of which is restricted to proliferating neuroectodermal cells, thereby providing a marker for neural differentiation (Ying et al., 2003b). Initially, two neural differentiation protocols were tested, whereby the first one led to the formation of embryoid bodies (EBs). When ESCs are grown to high density and in non-coated plates, they spontaneously form cell aggregates called embryoid bodies, which contain cells from the three germ layers (Bain *et al.*, 1995). Once formed, the embryoid bodies were then dissociated with trypsin and cultured as a monolayer in the absence of LIF. The neural cell morphology in the cell monolayer and the concomitant expression of Sox1:GFP confirmed that the neural differentiation occurred. However, the major limitation of this approach resides in the generation of an extremely heterogeneous cell population (data not shown). In fact the size of the EBs, strictly correlated to the initial cell density, has already been shown to greatly influence the lineage derivation efficiency (Zhou *et al.*, 2008).

In order to overcome this problem, a second protocol that comprises of monolayer differentiation of mESCs into neural cells was employed. As mentioned before, this protocol provides culturing ES cells in defined free serum condition and following LIF withdrawal (Ying et al., 2003b). In order to evaluate the efficiency of the *in vitro* neural differentiation, colonies were examined by fluorescent microscopy for GFP and images acquired over time (Figure 4.2).

It was extremely fascinating to observe how cell phenotype gradually changed during the neural differentiation. After three days, few colonies of mESCs started to express low level of Sox1:GFP that became stronger and more homogenous by day 6. The increase in Sox1:GFP signal was accompanied by changes in cell morphology whereby cells started to assume a rosette conformation (typical of neural epithelial precursors). By day 9, the neural precursors exhibited a clear and perfectly distinguishable neuronal morphology with extended protrusions that connects cells to each other.

The cell morphology clearly confirmed that the neural differentiation had successfully occurred but further validation at a molecular level came from the Western blot analysis (Figure 4.3). The protein levels of Nanog and Klf4 were dramatically reduced already two days after seeding, whereas protein levels of both Oct4 and Sox2 slightly increased during the first days of differentiation to be then gradually reduced. The reduction of these factors and the concomitant increase in Sox1:GFP level, which became detectable only towards the last days of differentiation, confirmed that the mESCs gradually lost their pluripotency features to simultaneously gain the neural marker expression. Interestingly, RNMT levels remained stable throughout the differentiation process, whereas RAM expression initially increased, to then dramatically drop to nearly undetectable levels. This result was found to be consistent amongst three independent biological replicates. These data perfectly reproduced the observation made with the mouse brain, which was characterised by high level of RNMT and low level of RAM expression.

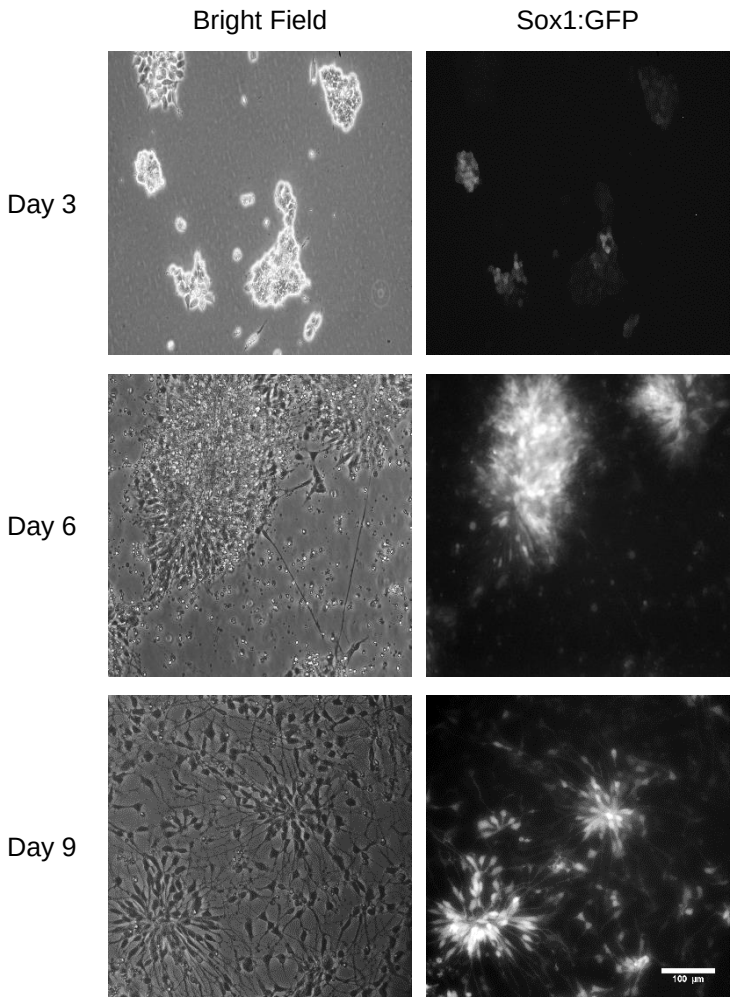


Figure 4.2: Morphology of mESCs changes during neural differentiation in N2B27 media.
mESCs were seeded into a 10 cm dish coated with 0.1% gelatin in N2B27 media for 9 days. Media was changed every second day. Images were captured on days 3, 6 and 9. Scale bar measures 100 µm. The above images are representative of six biological replicates.

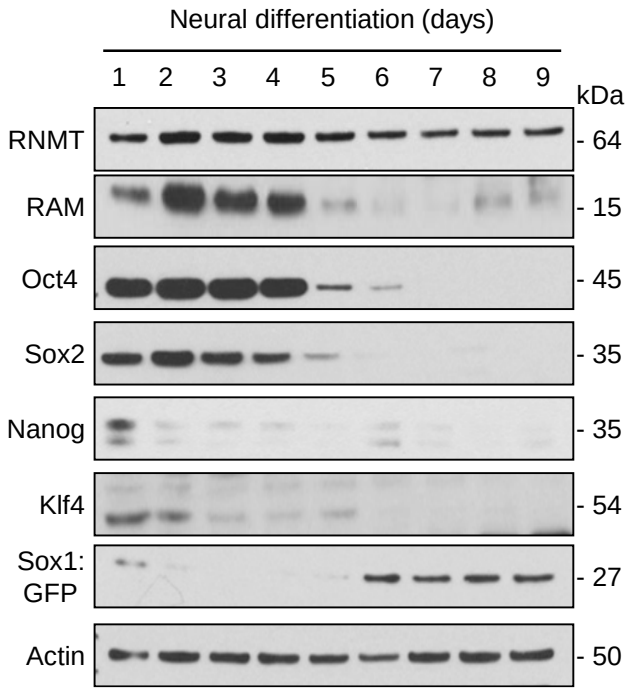


Figure 4.3: RAM expression is repressed during neural differentiation.
mESCs were seeded into a 10 cm dish coated with 0.1% gelatin in N2B27 media for 9 days. Media was changed every second day. Cell extracts collected at 24 hr intervals were analysed by Western blot with antibodies against the indicated proteins. Actin serves as loading control. The above image is representative of six biological replicates.

To investigate whether the reduction in RAM expression during *in vitro* neural differentiation was due to transcriptional down-regulation, RT-qPCR was performed (Figure 4.4). The mRNA levels of RNMT, Klf4 and Nanog were consistent with the protein expression with RNMT stably expressed and Klf4 and Nanog dramatically reduced after two days of neural differentiation. On the contrary, Sox2 transcripts differently from the protein levels that gradually decreased over time, halved on day 2 and remained so until the end of the protocol. Oct4 transcripts were gradually reduced over time. Interestingly, the mRNA levels of RAM did not significantly go down and remained stable over time whereas Nestin, neural progenitor marker (Lendahl *et al.*, 1990), with its gradual increase over time confirmed the emerging phenotype. Taken together these data suggest that RAM expression is not transcriptionally regulated during *in vitro* neural differentiation but other regulatory mechanisms at post-transcriptional level are possibly involved.

Validation of the neural differentiation came also from Immunofluorescence study where mESCs, seeded onto Laminin coated coverslip in N2B27 media, were fixed and stained on day 3, 6 and 9 (Figure 4.5). Immunofluorescence data were consistent with the analysis of protein levels, where the RNMT signal was invariant throughout the protocol (Figure 4.5a). High level of Oct4 expression were initially observed but then the signal intensity per nucleus decreased becoming weaker towards the last day. On day 9, those cells with low level of Oct4 showed a concomitant increase in the GFP signal (from Sox1:GFP) indicating that cells were differentiated (Figure 4.5). RAM signal, as previously observed by Western blot, decreased over time (Figure 4.5b). Noteworthy, on day 6, the few cells which retained Oct4 expression retained

high levels of RAM expression too. In addition, cells were also stained for Pax6, another neural marker (Walther & Gruss, 1991), and its gradual increase was perfectly consistent with the expression pattern of Sox1:GFP levels (Figure 4.5c).

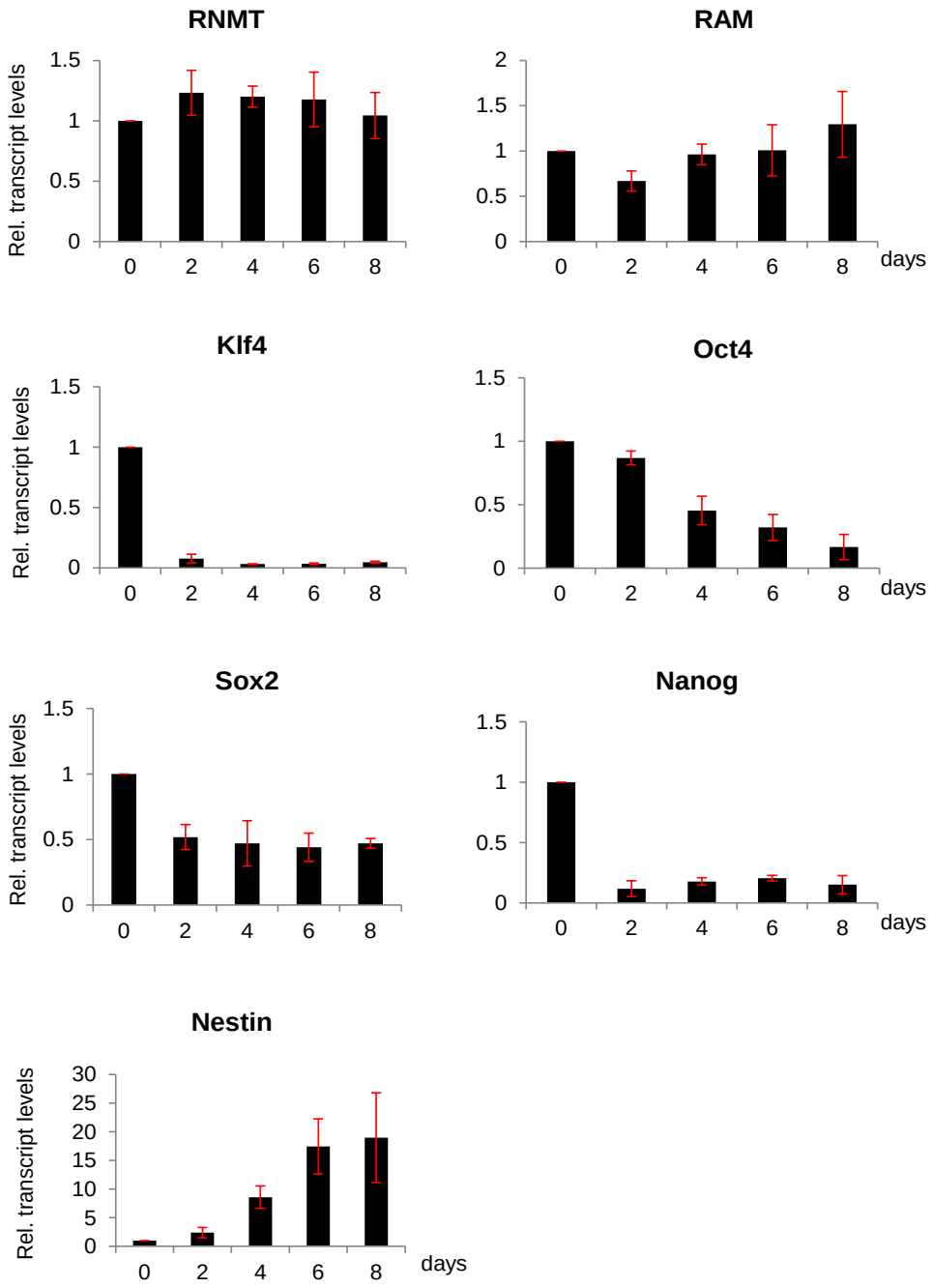
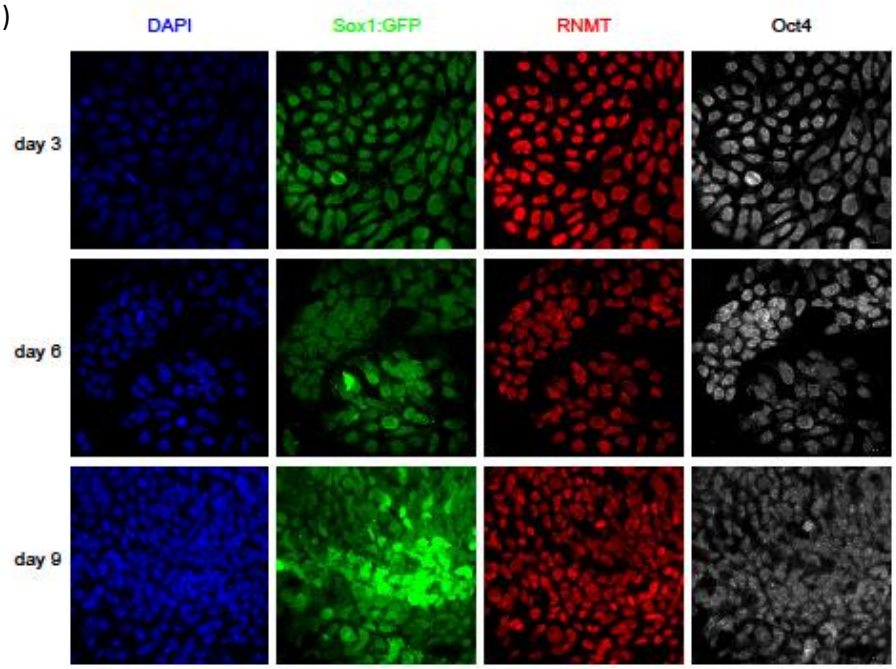
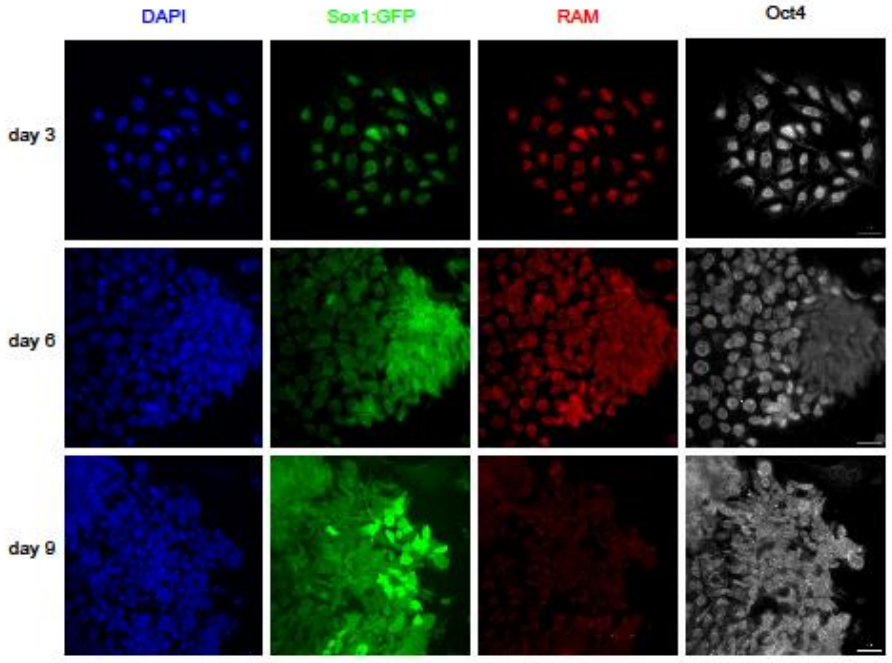


Figure 4.4: RAM expression during *in vitro* neural differentiation is not regulated at the level of the transcripts.
mESCs were seeded into a 10 cm dish coated with 0.1% gelatin in N2B27 media for 9 days. Media was changed every second day. RNA extracted on day 2, 4, 6 and 8 of differentiation was analysed by RT-qPCR. Day 0 refers to undifferentiated mESCs. The data were normalised to Actin mRNA levels. Bar charts depicts the average values and relative standard deviation of three independent experiments. The above images are representative of three biological replicates.

(a)


DAPI
Sox1:GFP
RNMT
Oct4
day 3
day 6
day 9

(b)


DAPI
Sox1:GFP
RAM
Oct4
day 3
day 6
day 9

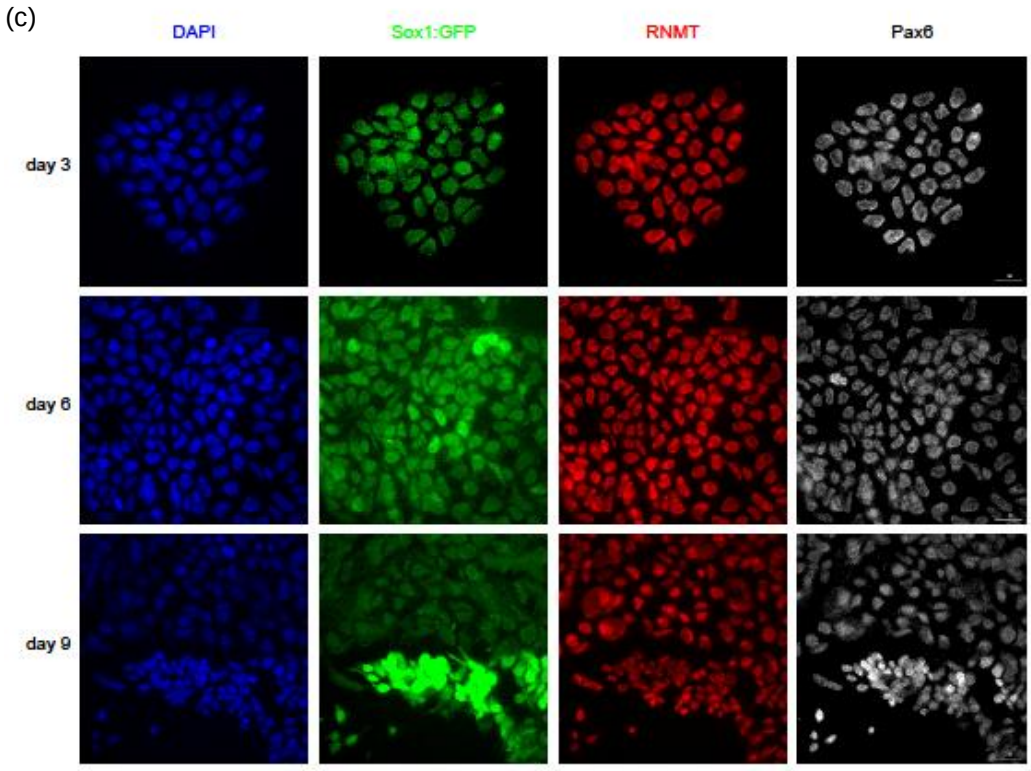


Figure 4.5: Immunofluorescence confirms the neural differentiation of mESCs and the associated reduction in RAM expression.
Immunofluorescence microscopy was used to confirm the neuronal differentiation. mESCs were seeded onto coverslip coated with Laminin in N2B27 media. Cells were fixed with 4% PFA and stained with the indicated antibodies on days 3, 6 and 9. DAPI staining was used to detect nuclei. High-resolution images were collected with an imaging system (DeltaVision Restoration; Applied Precision) using a 40X/1.514 oil (Olympus) objective lens. Images were then processed using OMERO software. Scale bar measures 20µm. (a) RNMT staining; (b) RAM staining; (c) Pax6 staining. The above images are representative of two biological replicates.

Considering the data so far, it was interesting to further investigate how the RNA methyltransferase activity was affected during *in vitro* neural differentiation. In order to address this, the *in vitro* methyltransferase activity assay was performed (Figure 4.6). From titrations performed in Figure 3.4, 0.5 µg of mESC protein lysate was the minimal amount of protein required for maximal conversion from cap to methyl cap. Thus, the subsequent assay was performed with 0.5 µg lysate of mESCs and *in vitro* differentiated neural cell. Cells were differentiated in N2B27 media as previously described, and according to the Western blot analysis, cells on day 8 were found to contain minimal RAM levels (Figure 4.6a). Recent unpublished data have shown that RAM may help recruitment of the S-adenosylmethionine (SAM), the methyl donor, to RNMT (data not show, Dr. Dhaval Varshney). Thus, to test whether the low level of RAM on day 8 of neural differentiation may affect SAM recruitment and thus the methyltransferase reaction, the *in vitro* MT assay was performed for 10 min in presence and absence of SAM (Figure 4.6b). Surprisingly, in both cases, the methyltransferase activity of neural cell extracts was found to be slightly higher than mESC extracts. The presence or absence of SAM did not alter the percentage methylation indicating that high levels of SAM are present in cell extracts. To remove SAM from cell extracts different lysis buffers were tested but they were found to inefficiently extracting the cell proteins (data not shown). Taken together these data suggest that neural differentiated cells, despite containing low levels of RAM, still have high levels of methyltransferase activity.

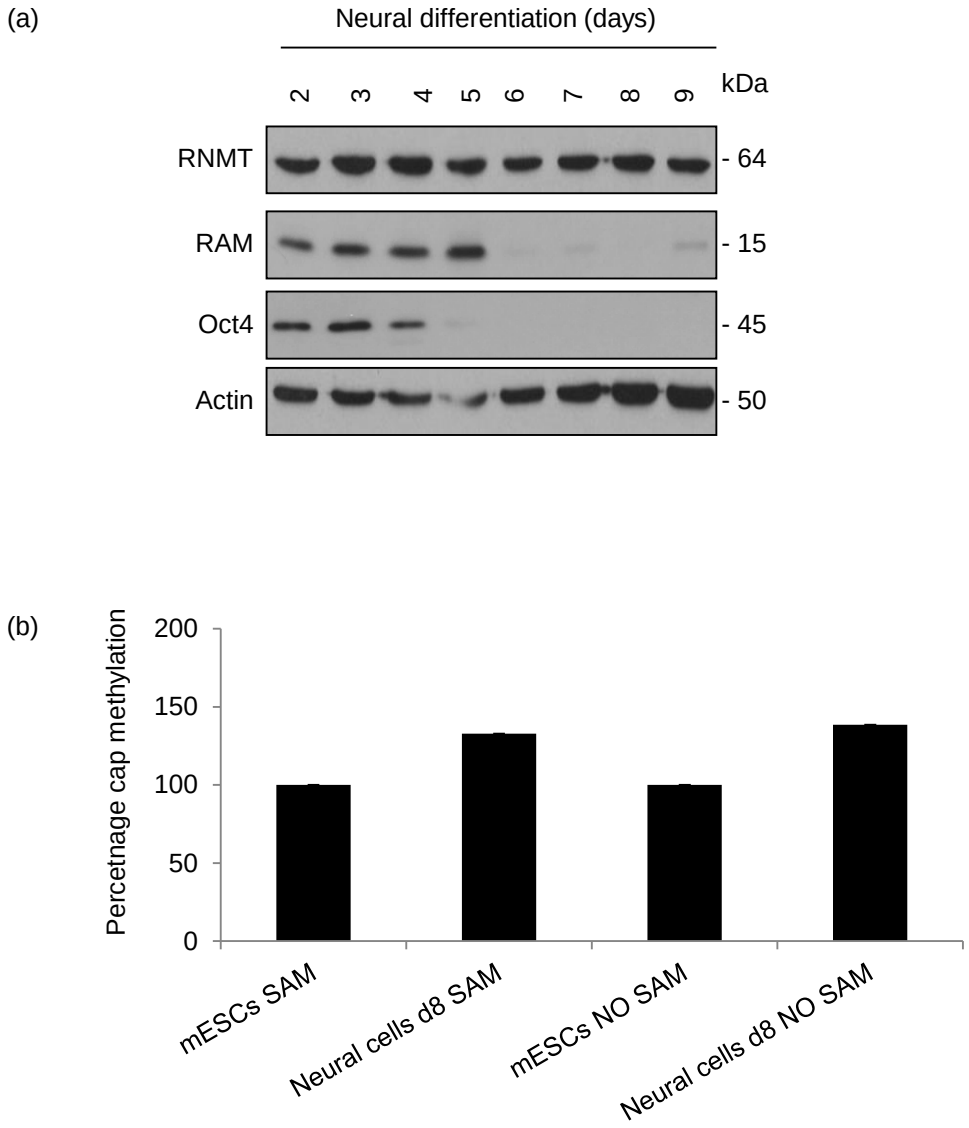


Figure 4.6: RNMT activity is slightly higher in neural cells than mESCs.
(a) mESCs were seeded into a 10 cm dish coated with 0.1% gelatin in N2B27 media for 9 days. Media was changed every second day. Cell extracts collected at 24 hr intervals were analysed by Western blot with antibodies against the indicated proteins. Actin serves as loading control. (b) Relative cap methyltransferase activity was detected in cell extracts using an *in vitro* cap methyltransferase assay. Activity was measured for 10 min with 0.5 µg of mESCs lysate and the neural cell. Quantification of thin layer chromatography was performed using ImageJ densitometry and the activity expressed as percentage methylation (i.e. (m^7GpppG x 100) / (m^7GpppG + GpppG)). The above data are representative of two biological replicates.

4.2.3 *In vitro* neural differentiation with hESCs

Lastly, the *in vitro* neural differentiation was repeated using hESCs. Dr Lindsay Davidson kindly provided and seeded the cells for this experiment. The protocol was performed for 13 days and on each day cells were collected for analysing protein and mRNA levels (Figure 4.7a and b). Excitingly, the neural differentiation with hESCs confirmed the expression pattern of RNMT and RAM observed with mESCs, whereby RNMT protein expression is maintained over time and RAM is similarly lost after day 6. The progressive decrease in protein levels of OCT4 and SOX2 confirmed the loss of pluripotency. Opposite trend was observed for PAX6, human neuroectodermal marker as well (Zhang *et al.*, 2010), which was detectable only towards the end of the protocol (the day one signal may be caused by blot dirtiness), confirming that the neural differentiation had occurred.

Consistent results were also obtained with RT-qPCR, where RNMT and RAM mRNA levels behaved similarly to what previously observed in mESCs as they remained unaltered over the differentiation protocol (Figure 4.7b). OCT4 mRNA expression decreased over time, whereas the mRNA levels of SOX2, despite significantly fluctuating, never decreased. PAX6 expression increased over time consistently to the protein levels. Moreover, Dr. Lindsay Davidson fixed and stained hESCs neural differentiated to confirm the differentiation process (Figure 4.7c). Immunofluorescence was performed by on day 11 and revealed high expression levels of two neural markers, Tubulin-βIII and PAX6. Due to limited sample availability it was not possible to perform *in vitro* methyltransferase assay with hESCs and the derived neurons.

The data indicate that the regulation of RNMT and RAM levels during neural differentiation appears to be conserved between mouse and human ESCs, where RNMT expression is maintained and RAM strongly down-regulated.

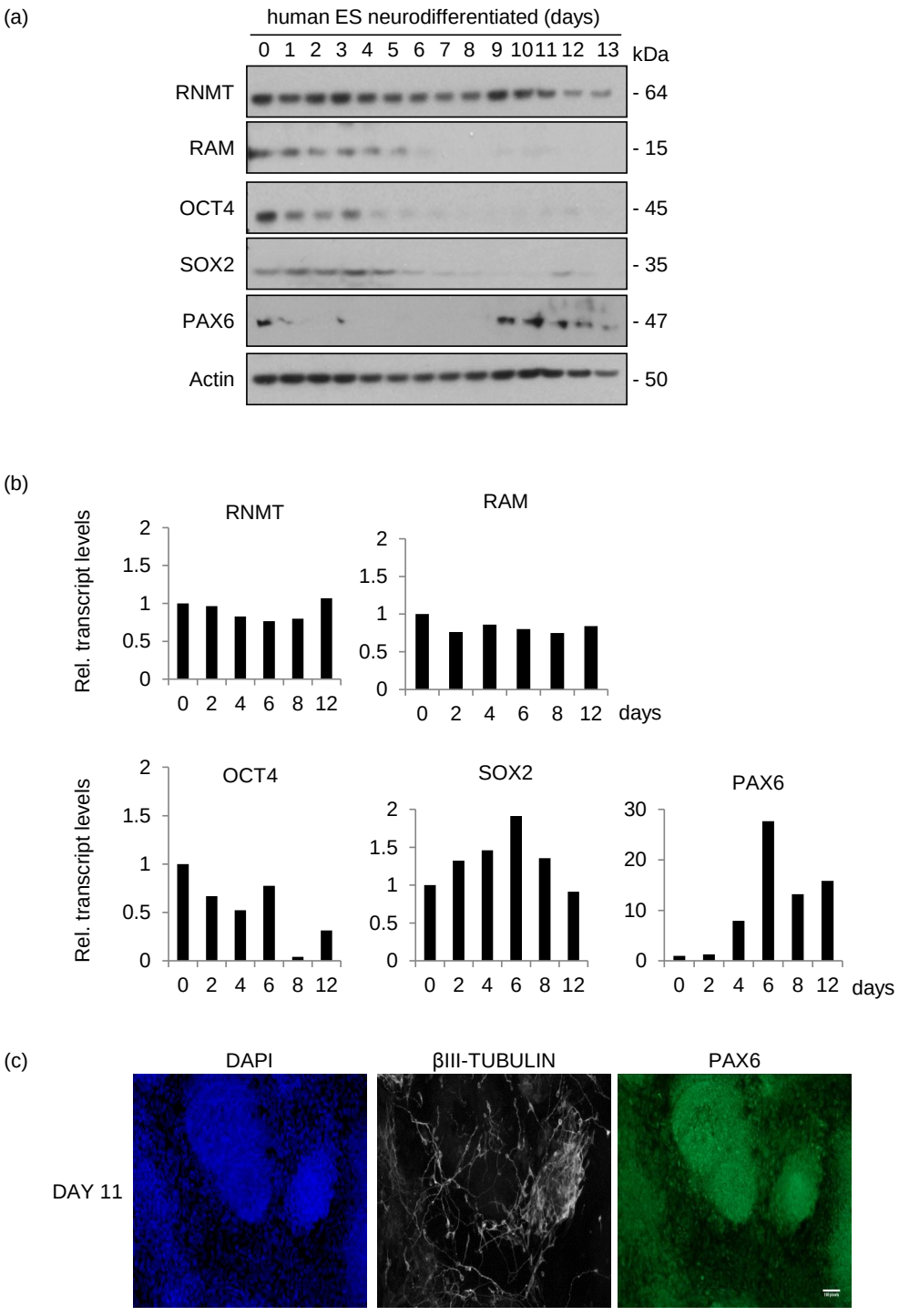


Figure 4.7: RAM expression is also reduced upon neuronal differentiation of hESCs.
(a) Cell extracts were collected at 24 hr intervals of the neural differentiation and analysed by Western blot with antibodies against the indicated proteins. Actin serves as loading control. (B) RNA extracted on day 0, 2, 4, 6 and 8 of the differentiation process was analysed by RT-qPCR. The data were normalised to GAPDH mRNA levels. (C) Immunofluorescence of human ES cells differentiated on day 11 was performed by Dr. Lindsay Davidson (Cell and Developmental Biology, University of Dundee). Cells were fixed with 10% neutral buffer formalin and stained with the indicated antibodies. Nuclei were stained with DAPI. Scale bar measures 100 µm. The above images are representative of two biological sample.

4.3 Discussion

In the previous chapter, RNMT and RAM were found to be highly expressed in pluripotent stem cells when compared to differentiated cells. In addition to this, it was shown that in different mouse organs, the stoichiometry of RNMT and RAM expression varies according to the cell lineage commitment. For example, the brain represents a really interesting case-study as RNMT is highly expressed whereas RAM is surprisingly not. With the purpose to better understand the dynamics of RNMT and RAM regulation, we decided to *in vitro* differentiate mESCs into neural lineage commitment and monitor the expression of the two proteins. The result was extremely interesting because it mirrors the data obtained with the brain; in fact, during *in vitro* neural differentiation high levels of RNMT expression were maintained whereas RAM was dramatically reduced. This observation is important because it highlights the reproducibility between i*n vivo* data (mouse organs) with *in vitro* results (*in vitro* neural differentiation).

In order to elucidate whether RAM down regulation occurs at transcriptional level, RT-qPCR was performed. Data revealed that the protein was not transcriptionally regulated, highlighting a certain discrepancy between transcript and protein levels. However, it is important to stress that RT-qPCR only measures the amount of viable mRNAs and regulation of transcription represents only one of the multiple mechanisms by which the level of the final protein is controlled. In fact, decrease in the gene product may be caused by problems related to the incorrect folding of the proteins or to lysosomal degradation.

Moreover, a central role is played by covalent modifications described as post translation modifications, e.g. addition of ubiquitin labels proteins for proteosomal degradation (Chau *et al.*, 1989). PTMs may also interfere or favour the association of the protein with others, leading to activate or inactivate the protein complex. Possible mechanisms of RAM regulation will be discussed in chapter 6 of this thesis.

Previous findings in our laboratory have shown that RAM stimulates RNMT activity both *in vitro* and *in vivo* (Gonatopoulos-Pournatzis et al., 2011). *In vitro* RAM has been found to highly promote recombinant RNMT activity and in cellular extracts RAM depletion reduced the methyl cap levels on endogenous transcripts. Considering these evidence we next measured the methyltransferase activity to see whether the down-regulation of RAM in neural cells affects the ability to methylate capped mRNAs. Unexpectedly, the methyltransferase activity was slightly higher in neural cell lysates compared to mESCs. One hypothesis is that due to the different nucleus/cytoplasm ratio between mESCs and neural differentiated cells, during the protein extraction other enzymatic activities e.g. decapping enzymes, could have interfered with the assay. To discard this possibility, mESCs expressing HA-RNMT were generated and in future will be neural differentiated, RNMT will be immunoprecipitated and the assay will be repeated. However, this result combined with the previous gel filtration and IP data (not all RNMT may be bound to RAM in mESCs) could be explained by two hypotheses: the first being that RNMT may interact with other RAM-like proteins that somehow replace RAM during lineage specification.

The second explanation could be that monomeric RNMT in stem cells does not need RAM to increase its already high activity. Furthermore, if the preliminary data obtained in our laboratory about RAM facilitating the recruitment of SAM to RNMT (data not published, Dr Dhaval Varshney) will be validated, in agreement with the second scenario, another possible explanation of higher MT activity in neural cells is that those cells have excess of SAM and therefore RAM is not essential. However, this latter hypothesis does not consider the evidence that RAM contains the RNA binding domain within the RNMT complex and therefore RAM has the role to bind RNA and make it available for the cap methylation. According to this, changes in RAM expression may be explained by speculating that RAM does not bind all mRNAs but specifically recognizes certain transcripts required for that particular cell state. Thus, relatively to the transcriptional changes that occur during stem cells differentiation, RAM may be less or more required e.g. respectively brain and heart. Taken together these observations suggest that RAM expression may be context–dependent and emphasise the potential role of cap methylation in stem cell fate determination.

Extremely exciting are the *in vitro* neural differentiation data with human ESCs. In fact, as it has been reported for the mESCs, the *in vitro* neural differentiation is able to derive neural progenitors in a rosette-like structure that recapitulates the neural tube formation (Gerrard *et al.*, 2005). These evidence strongly strengthen the reliability of the *in vitro* observations because the same dynamics will very likely occur *in vivo*. Data obtained with hESCs perfectly reproduced the same pattern of RNMT and RAM expression described in mESCs indicating that the mechanisms regulating RAM expression appear to be conserved in mice and humans. The *in vitro*-neural differentiation has been

repeated with H9 hESCs and the lost in RAM expression was confirmed (data not shown). Unfortunately, MT activity was not performed with *in vitro* neural differentiated hESCs and it will be interesting in the future to repeat the assay to see whether the same results are obtained. These evidences strongly encourage further investigations particularly into the role of RAM in stem cell fate decision, and raise the question whether RAM reduction during neural differentiation may be a cause or a consequence of the lineage specification process.

5 Investigating the biological role of RAM in embryonic stem cells

5.1 Introduction

Eukaryotic protein synthesis comprises three stages and the rate-limiting step is usually translation initiation. In the cap-dependent translation, eIF4F complex via eIF4E, its cap binding subunit that specifically recognizes the methyl cap at the 5' end of mRNA, mediates the interaction between mRNA and 43S pre-initiation complex (PIC) (Sonenberg & Hinnebusch, 2009). Once the PIC is bound to the cap proximal region of mRNA, it starts to scan the 5'-untranslated region in the 5'-3' direction until if finds the start initiation codon. As soon as the AUG codon is found, a dynamic change and removal of factors results in the formation of elongation-competent 80S ribosome (Jackson et al., 2010).

Polysome profiling relies on the ability of mRNA to recruit more than one 80S ribosome resulting in the so called polysome. Therefore the polysome profiling analysis is a reliable method to characterize mRNA actively engaged in translation (Masek *et al.*, 2011). However, because translation is a highly dynamic process, cells are treated with the translational elongation inhibitor cyclohexamide in order to prevent polysome run–off during sample preparation. Cyclohexamide is the most common inhibitor of eukaryotic translation and it functions by binding 60S and blocking the release of the deacylated tRNA from ribosome E site, thus stalling the ribosome in a polysome state (Schneider-Poetsch *et al.*, 2010). Polysome preparation can then be separated through a sucrose gradient thus allowing the subsequent fractionation of mRNAs according to the number of ribosomes bound. This technique divides mRNAs into: monosomal (free ribosomal subunits and mRNA associated with few ribosome) and polysomal (mRNA bound at multiple ribosome) fractions.

In the previous chapters, RNMT and RAM were shown to be highly expressed in ESCs, whereas *in vitro neural* differentiated cells were found to contain dramatically reduced levels of RAM. These observations suggest that high levels of RNMT and particularly of RAM, may be required to maintain pluripotency. Therefore, this chapter will investigate the biological role of RAM in mESCs.

5.2 Results

5.2.1 Depletion of RAM impairs Oct4 expression

In order to address the aim of this chapter, a loss of function study was performed using the RNAi methodology. To rule out off-target effects and considering that both RNMT and RAM are highly expressed in mESCs and are both required for optimal methylation of the cap, the majority of the knockdown assays were carried out with three independent siRNAs. The efficiency of the siRNAs has already been tested in chapter 3, whereby the siRNAs were employed to verify the specificity of RNMT and RAM antibodies. Briefly, cells were seeded and immediately transfected with siRNAs and with non-targeting control. To test the knockdown efficiency, cells were lysed 2, 3 and 4 days following transfection. Initial experiments indicated that two days of treatment with siRNAs was sufficient to obtain a substantial depletion of the proteins (data not shown), thus all further experiments described in this chapter were performed with a 48 hr siRNAs treatment. The Western blot depicted in figure 3.2, showed that all the siRNAs employed successfully impaired the expression of their target protein compared to cells incubated with non-targeting control.

Quantification of the western blot (Fig 5.1c), showed that depletion of RNMT or RAM had similar effect on the other protein levels, with particularly the siRNA number 3 affecting the expression of the interacting partner. In order to evaluate the impact of RNMT or RAM knockdown on cell pluripotency, Oct4 levels were investigated (Figure 5.1a). Interestingly, all three siRNAs against RAM consistently impaired Oct4 expression, whereas RNMT knockdown did not lead to consistent alterations as siRNA 3 caused reduction in Oct4 levels and siRNA 2 caused an increase. This data represents the first evidence that directly correlates RAM expression with one of the pluripotency factors.

5.2.2 High RAM expression is required for the maintenance of pluripotency

The finding that RAM alters Oct4 expression combined with the previous data strongly supports the hypothesis that high levels of RAM specifically correlate with a pluripotent state and prompted us to focus on the role of RAM in mESCs. To further investigate how RAM depletion affects stem cell, its impact on cell proliferation was determined (Figure 5.1c). None of the three siRNAs caused significant delay in cell growth as cells transfected with siRNA showed similar level of cell proliferation when compared to control transfected cells.

Next, the ability to form undifferentiated colonies following RAM depletion was assessed. To do this, cells were dissociated following 48hr siRNA treatment, re-seeded at a low density and allowed to grow for 7 days in the usual mESCs media. Cells where then fixed and stained for AP activity (Figure 5.2a). According to the colony morphologies, the associated intensity and distribution of the AP staining, three different populations were identified: undifferentiated

colonies were characterised by homogenous edges and intense purple staining; mixed colonies were poorly stained and with deformed edges (sign of differentiation) and lastly differentiated colonies, which lacked clear staining (Fig 5.2a). Images of the different colonies were taken from cells treated with non targeting control and were representative of the morphology observed in experimental conditions too. The percentage of undifferentiated, mixed and differentiated colonies was calculated for each transfection condition (Figure 5.2b). It emerged that RAM depletion impaired the ability of cells to form undifferentiated colonies. More precisely, following siRNA 1 treatment against RAM, cells showed a slight reduction in the number of positive colonies, on the contrary when cells were treated with siRNA 2 and 3 against RAM, the number of mixed colonies increased significantly compared to the control. These data suggest that despite not being essential for cell proliferation, high level of RAM in stem cells may be required for the maintenance of pluripotency. Its depletion reduces Oct4 expression and may therefore push cells towards differentiation.

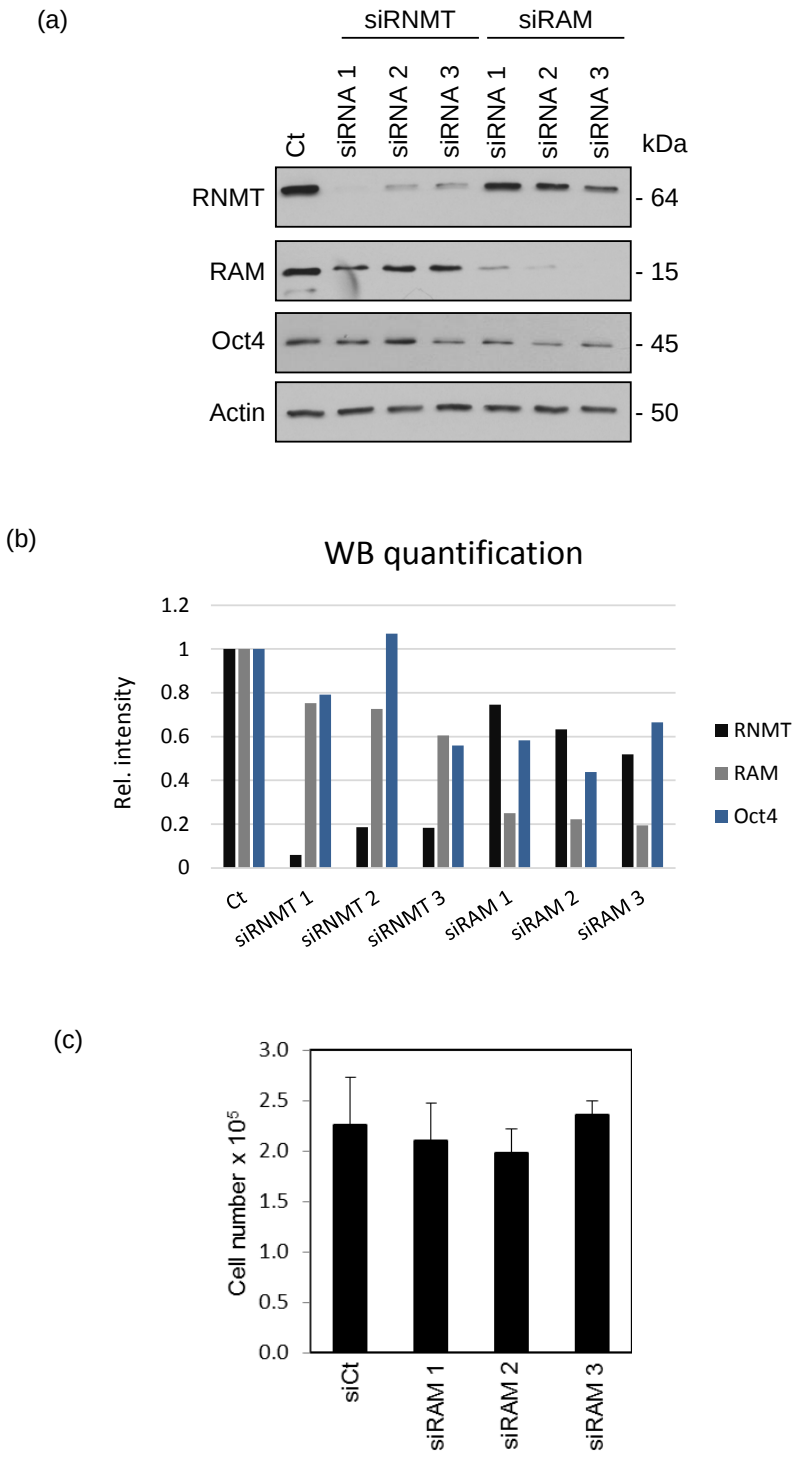
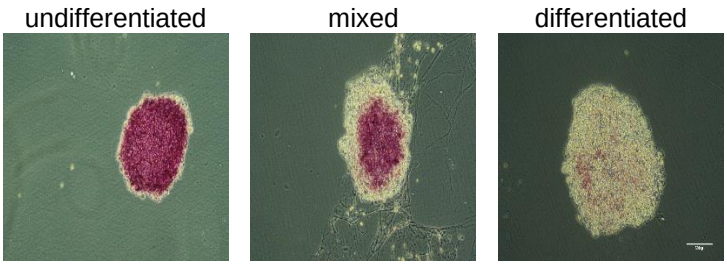


Figure 5.1: RAM knockdown has an effect on Oct4 expression but does not effect cell proliferation.

mESCs were transfected with three independent siRNAs against RNMT, three independent siRNAs against RAM (1, 2, and 3) and a non-targeting control (siCt), for 48 hr. (a) Cell extracts were analysed by Western blot with antibodies against the indicated proteins. Actin serves as loading control. (b) Quantification of western blot was performed with Image J Software. (c) Cell counts were determined 48 hr post-transfection using a Countess cell counter. Bar charts depict the average cell number and relative standard deviation of three independent experiments (n=3).

(a)



(b)

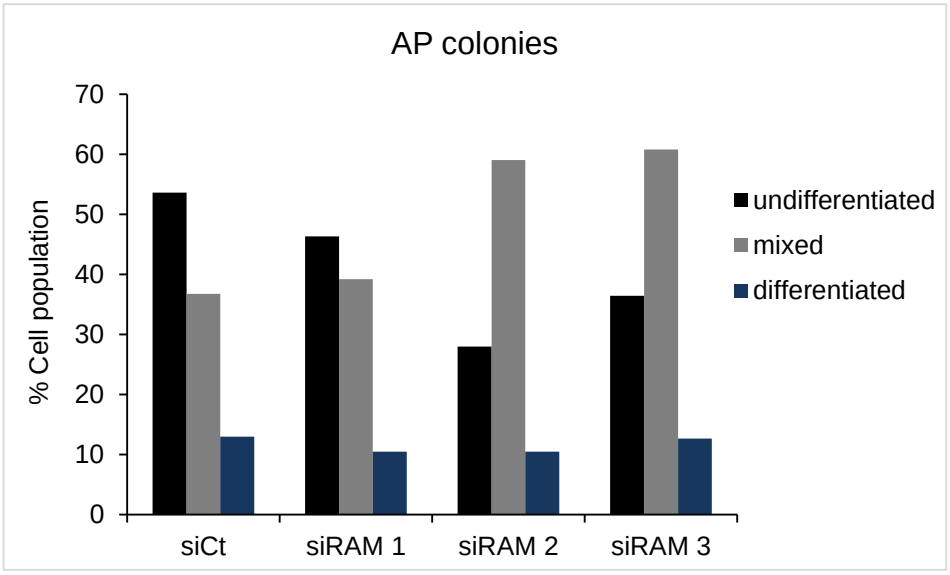


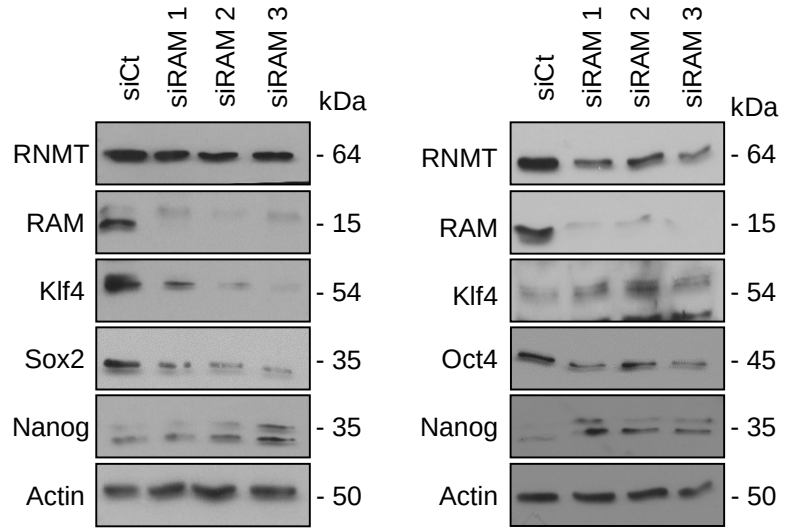
Figure 5.2: RAM knockdown alters the ability of mESCs to form undifferentiated colonies.

mESCs were transfected with three independent siRNAs (1, 2 and 3) against RAM or with non-targeting control (siCt) for 48 hr following which cells were dissociated with trypsin and plated in a 6 well plate. After one week colonies were stained for Alkaline phosphatase activity. Colonies were categorized as undifferentiated (stained with intense purple), mixed (poorly stained and with deformed edges) and differentiated (not stained). (a) Bright field image of colonies of cells treated with siRNA control and classified as undifferentiated, mixed and differentiated. Images were taken with Canon camera EOS 100D. Scale bar 500 µm. (b) Graphs display the percentage of undifferentiated, mixed or differentiated colonies for the siCt and for each individual siRAM. The percentage was calculated on the average value of three wells. The above images are representative of two biological replicates.

5.2.3 RAM knockdown affects cap methylation of certain transcripts

In order to gain further insight into how depletion of RAM impacts pluripotency, Western blot and RT-qPCR analyses of the major stemness markers Oct4, Sox2, Nanog and Klf4 was performed. Western blot analysis was carried out more than 5 times and the most representative results, chosen according to the quality of the blots, are presented (Figure 5.3a). As expected, RAM depletion caused alteration in the protein levels of some of the pluripotency factors. In particular, Oct4, Sox2 and Klf4 levels in every experiment were found to decrease more consistently with all the three siRNAs against RAM. Contrary to this, discordant results were obtained with Nanog, whereby protein levels were down regulated in 50% of the experiments and upregulated in the other half and the effect was not always consistent among the three siRNAs employed. The quantification of more than three western blot analyses corroborated as described above with RAM depletion having a major effect on three out four pluripotency factors analysed.

(a)



(b)

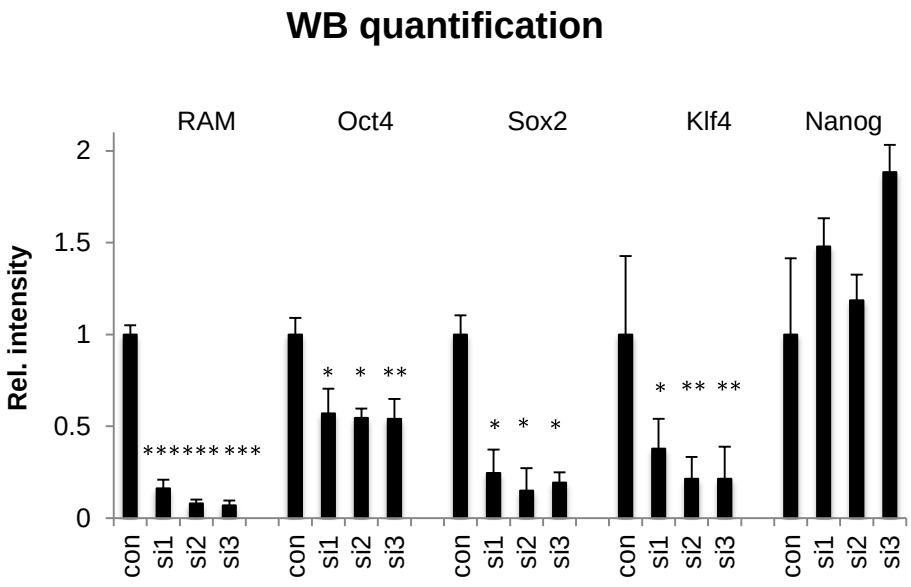


Figure 5.3: RAM knockdown affects Oct4 and Sox2 expression levels.

(a) Western blot analysis with antibodies against the indicated proteins on extracts from mESCs 48 hr following transfections with three independent siRNAs (1, 2 and 3) against RAM or with non-targeting siRNA (siCt). Actin serves as a loading control. The above images are representative of three biological replicates. (b) Quantification of western blot analyses was performed using Image J software. Bar charts depicts the average values and relative standard deviation of at least three biological replicates. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). * indicates $p<0.05$, ** indicates $p<0.005$ and *** indicates $p<0.001$.

Next, to assess whether this effects was mirrored at transcript level, RT-qPCR analysis was performed (Figure 5.4). The dramatic reduction of RAM mRNA levels clearly confirmed the knockdown efficiency and as expected, RNMT transcripts were not affected. If anything the depletion of RAM induced a slight increase in the transcript levels of RNMT. Surprisingly, silencing of RAM did not consistently affect the transcripts of the pluripotency factors as Oct4, Klf4 and Nanog mRNA levels were particularly downregulated with only siRNA number 2 whereas no impairment was revealed for Sox2 transcripts. Discrepancy between protein and transcript levels following RAM depletion implicates a post-transcriptional mechanism in the regulation of gene expression by RAM depletion.

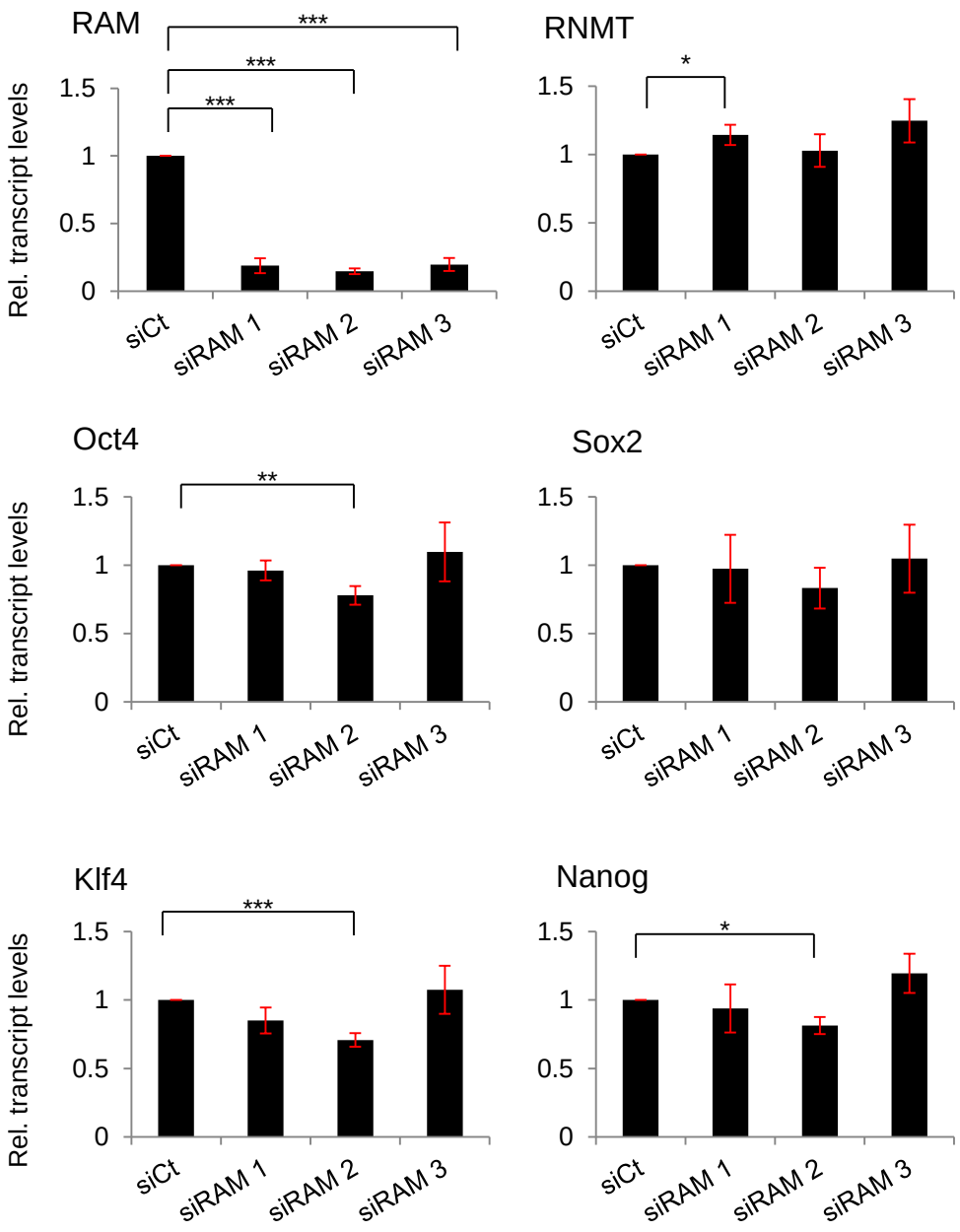


Figure 5.4: RAM knockdown does not affect the transcript levels of pluripotency markers.
RT-qPCR analysis performed on RNA from mESCs 48 hr post-transfection with three independent siRNAs (1, 2 and 3) against RAM or with non-targeting siRNA (siCt). The data were normalised to Actin mRNA levels. Bar charts depicts the average value and relative standard deviation of three independent experiments. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). * indicates $p<0.05$, ** indicates $p<0.005$ and *** indicates $p<0.001$. The above images are representative of three biological replicates.

As discussed previously, one of the multiple functions of the methyl cap is to recruit eIF4E and other translation initiation factors, thus promoting the engagement of the ribosome with mRNA, which leads to translation initiation (Topisirovic et al., 2011). To study whether RAM depletion could affect the ribosomal loading onto mRNA and thus explaining the reduction in the protein levels of the pluripotency factors, polysome profiling was performed. As mentioned earlier, this technique separates mRNA by ultracentrifugation across a sucrose gradient, fractionating mRNA into monosomal and polysomal fractions. Considering that the experiment needs to be performed quickly in order to avoid ribosomal run-off from the transcripts, and the high number of cells required, we decided to perform the analysis with control transfected cells or cells transfected with one siRNA against RAM. Given the Western blot data and particularly the AP staining where siRNA 2 against RAM caused a significant increase in the population of mixed colonies, siRNA 2 was chosen for the analysis (Figure5.5).

Following 48 hr siRNA treatment, cells were provided with fresh media in order to promote translation prior to cyclohexamide treatment. The RNA content and separation between monosome and polysome was monitored by UV absorbance at 254 nm and plotted (Figure 5.5a). The lighter monosomal fractions are found on the left side representing the top of the gradient, whereas the heavier polysome are found on the right side representing the bottom of the gradient. From the UV profiles, it is apparent that the siRNA against RAM caused a reduction in the overall translation rate of the cells as seen by the narrower and flatter polysome peak compared to the control.

After fractionation, RNA was extracted from each fraction and analysed by RT-qPCR (Figure 5.5b). Following RAM knockdown, mRNA distribution of Oct4, Sox2, Klf4 and Nanog, clearly showed a shift from the polysome (from fractions 6 to 8) to the monosome (from fractions 1 to 5) fraction. This effect seems specific since Actin mRNA was not found to follow this trend. The RT-qPCR of three biological replicates are further summarised in Figure 5.5c, where the RNA content of the polysomal fractions was pooled together for the siRNA against RAM treatment and for the control treated cells. Clearly, following RAM depletion, mRNAs expressing the pluripotency markers showed a substantial reduction in the amount of polysome engaged compared to the control treated cells. In fact, following RAM knockdown the amount of polysome loaded into the pluripotency factors Nanog, Klf4, Sox2 and Oct4 mRNAs was almost halved compared to the control cells. On the contrary, translation from Actin transcripts was not affected, indicating that the effect of RAM impairment is specific and not generalised to translation from all RNA Pol II transcripts. The diminished levels of polysome loading on transcripts of the pluripotency factors correlates with the reduction of their relative protein levels.

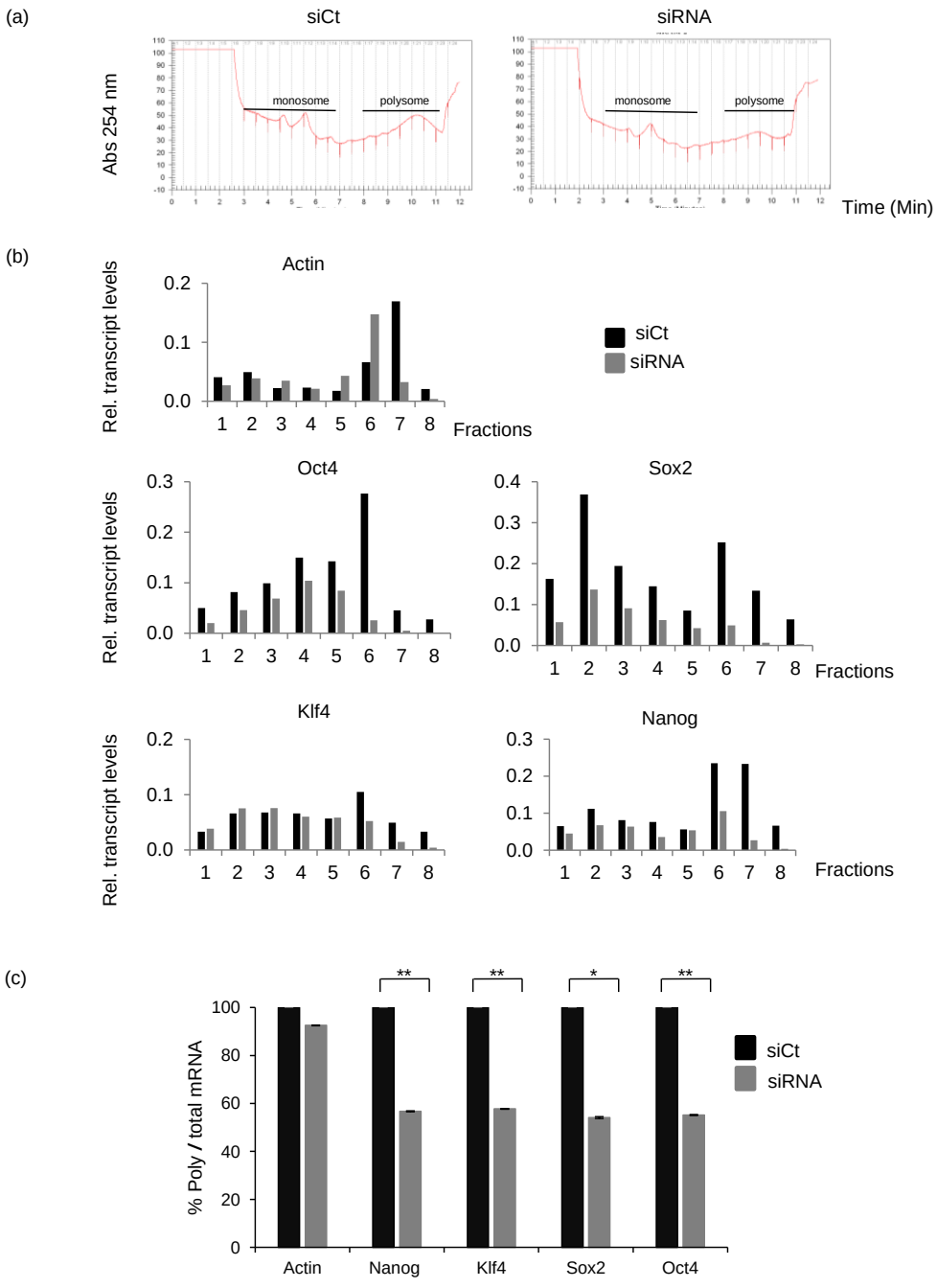


Figure 5.5: RAM depletion leads to reduced translation of the major pluripotency markers.

mESCs were transfected with siRNA 2 against RAM or non-targeting control and cells were collected 48 hr after transfection. Polysome associated RNAs were fractionated by sucrose-gradient ultracentrifugation. (a) Relative absorbance (254 nm) trace measuring the RNA distribution following polysome fractionation for siRNA or non-targeting control treated cells. (b) RT-qPCR analysis performed on fractions with primers against indicated target transcripts measuring their relative distribution amongst the fractions. (c) Percentage of a given mRNA present in polysomal fractions (6-8) to the total mRNA in all fractions (Polysome (6-8)/ total mRNA (monosome+polysome)). Ct values were normalised to the input (RNA prior fractionation). The graph is representative of three biological replicates. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). * indicates $p<0.05$, ** indicates $p<0.005$.

In addition to this, Dr. Victoria Cowling performed immunoprecipitation experiments using an antibody that specifically recognises methyl capped transcripts in order to investigate the effect of RAM knockdown on methyl cap formation on mRNAs expressing the pluripotency markers (Cole & Cowling, 2009) (Figure 5.6). It was observed that following RAM depletion with the three siRNAs, the amount of methyl capped Actin mRNA immunoprecipitated did not change compared to cells treated with non-targeting control as only siRNA1 slightly affected the m7G levels. On the contrary, following RAM depletion, the levels of methyl capped mRNAs for Oct4, Sox2 and Klf4 were clearly reduced with all the three siRNAs compared to the non-targeting control. Nanog mRNA was not efficiently immunoprecipitated in this experiment.

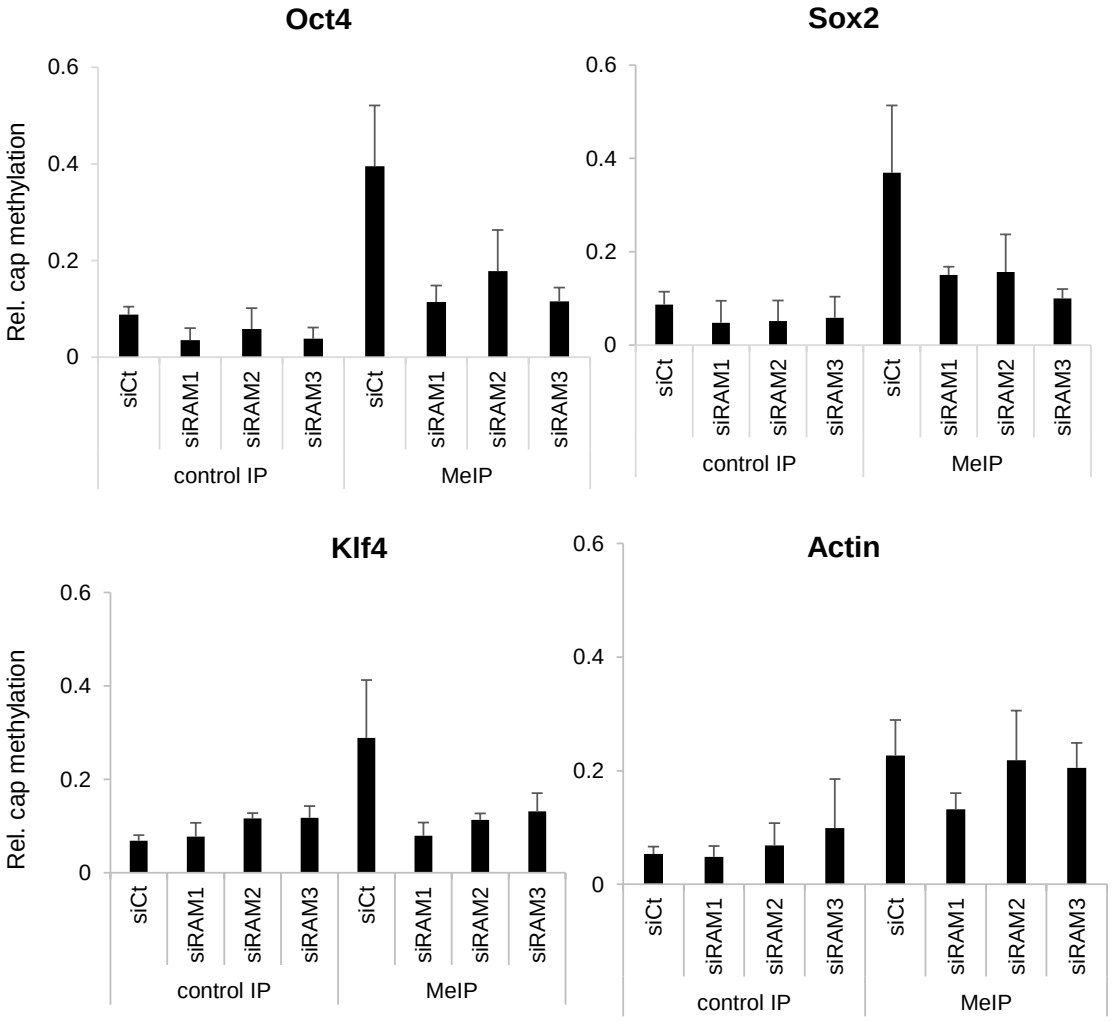


Figure 5.6: Depletion of RAM impairs the formation of the methylguanosine cap of pluripotency factors.
Data were kindly provided by Dr. Victoria Cowling. Methyl cap immunoprecipitations were performed on mESCs transfected with three independent siRNAs or non-targeting control. Methyl cap levels of endogenous transcripts indicated were determined relative to control IP using RT-qPCR following 7-methyl guanosine immunoprecipitations. Bar charts depicts the average values and mean relative standard deviation of three independent experiments. The above data are representative of three biological replicates.

Combined together, these results indicate that loss of RAM has some specificity towards the pluripotency factors transcripts that being efficiently cap methylated, are more efficiently engaged into the translation machinery thus explaining the elevated protein expression. On the other hand, following RAM impairment, reduced cap methylation of these transcripts occurs and consequently they are less efficiently translated resulting in lower protein expression. These data, for the first time, shed light on the function of RAM where it specifically recognises some transcripts rather than others. However, the exact mechanism for this selectivity is not clear and further studies will be needed to elucidate it.

5.2.4 RAM over-expression directly affects Oct4 and Sox2

So far, the loss of function studies have described a correlation between RAM and the pluripotency markers. To further confirm this relationship, gain function experiments where RAM was exogenously expressed were performed, using the *in vitro* neural differentiation model. As described previously, RAM protein levels were found to be down-regulated during the *in vitro* neural differentiation, therefore stable cell lines expressing exogenous RAM were created to study how the over expression of RAM impacts the differentiation process. Three different cells lines were made: mESCs transfected with empty vector, mESCs transfected with RAM fused to Green Fluorescent Protein on the C terminal domain (RAM-GFP) and mESCs transfected with RAM fused to Flag on the N terminal domain (FG-RAM). The expression of the exogenous protein was evaluated by Western blot analysis (Figure 5.7).

The double band in the RAM blot and the correspondent band in the Flag blot confirmed the expression of the FG-RAM exogenous protein (Figure 5.7a). For the RAM-GFP cell line, in the membrane blotted with GFP, the RAM-GFP expression was confirmed by the band of appropriate band (~ 40 kDa), the middle band at 35 kDa could be the product of an internal ribosome entry present on the GFP tag protein whereas the band at 27 kDa represented the GFP protein alone. Although the expression of GFP in ES cells derives from the expression of Sox1-GFP, the GFP protein was detectable also in ES. The explanation is likely due to the leaky GFP promoter (Stavridis, unpublished data). Expression of RAM-GFP was corroborated also by immunofluorescence (Figure 5.7c) that showed the RAM-GFP signal co-localise with the DAPI staining indicating that RAM-GFP was present in the nucleus of the cells.

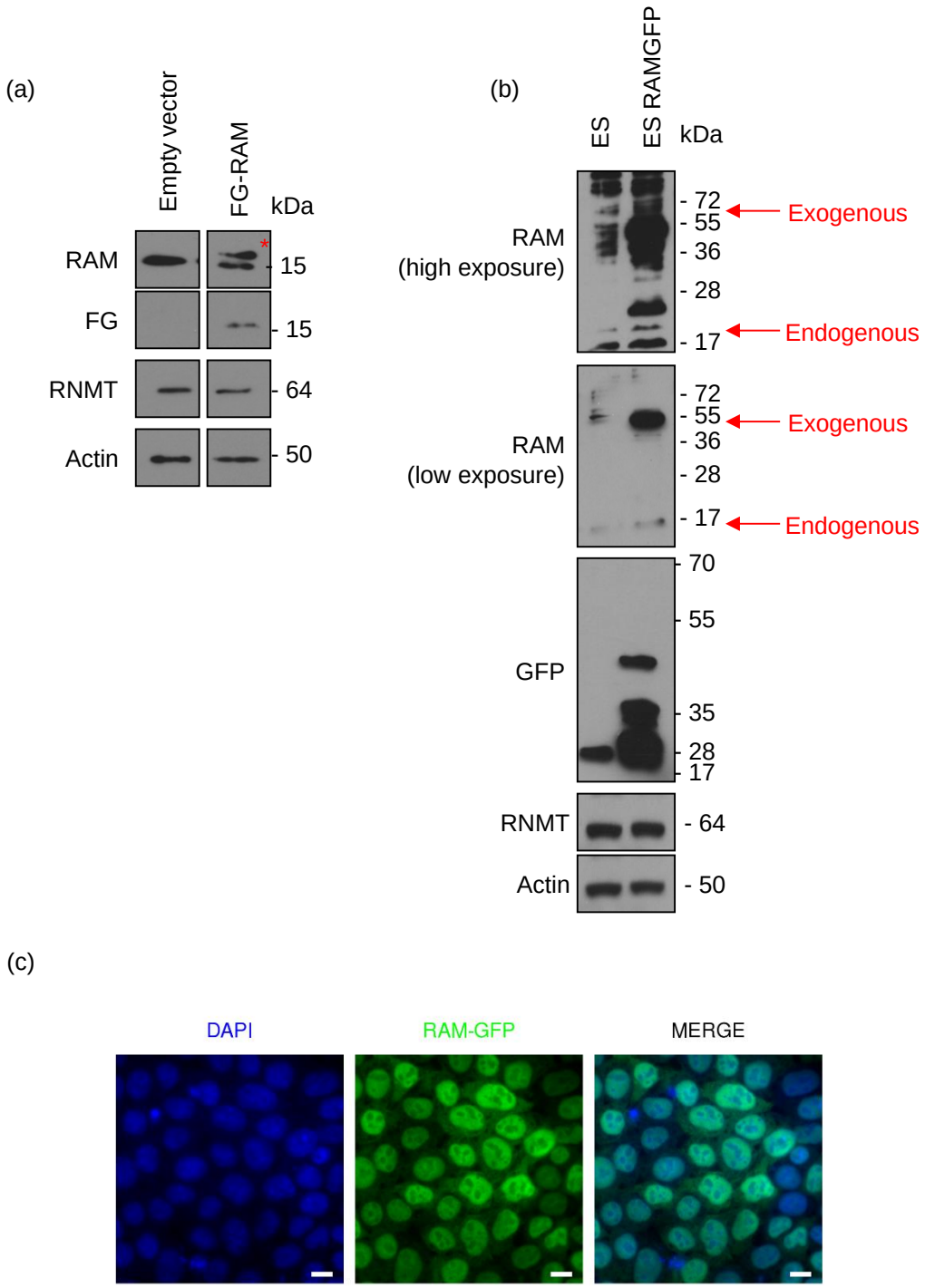


Figure 5.7: Characterization of the stable cell lines expressing exogenous RAM. Stable cells lines were created with mESCs transfected with pPyPCAGGS-Fg-RAM (FG), RAM-GFP (GFP) or with empty vector. (a and b) Western blot analysis on extracts from mESCs expressing exogenous RAM (FG or GFP) with the indicated antibodies. * indicates the FG-RAM band. Samples derive from the same gel. Actin serves as loading control. (c) Immunofluorescence of mESCs stably expressing RAM-GFP. Nuclei were stained with DAPI. Scale bar measures 10 µm.

A preliminary *in vitro* neural differentiation was performed with the three different cell lines in order to see whether the expression of exogenous RAM was sustained over the differentiation process or down-regulated similarly to the endogenous protein (Figure 5.8). Due to number of samples, the different cell lines were collected and analysed every three days. All the three cell lines successfully underwent neural differentiation as seen by the reduction in Oct4 expression by day 6. As expected RNMT expression was consistently maintained over time. In the cell line transfected with the empty vector, RAM was still detectable at day 6 but as previously observed disappeared at day 9. In the cell line transfected with FG-RAM, the levels of exogenous FG-RAM (the upper band in RAM blot labelled with *) were lower than the endogenous as early as day 3 and lost by day 6. Contrary to the exogenous RAM, endogenous RAM was barely detectable on day 6 but then reappeared again on day 9.

As opposed to FG-RAM, the RAM-GFP cells maintained high expression of the exogenous RAM (RAM-GFP detectable at the top of the blot) over time whereas the endogenous fluctuated as described for the FG-RAM cells. Given the more consistent and stable expression of exogenous RAM-GFP, this cell line could be used as system where RAM expression is maintained over the differentiation process and thus study its effect on pluripotency factors.

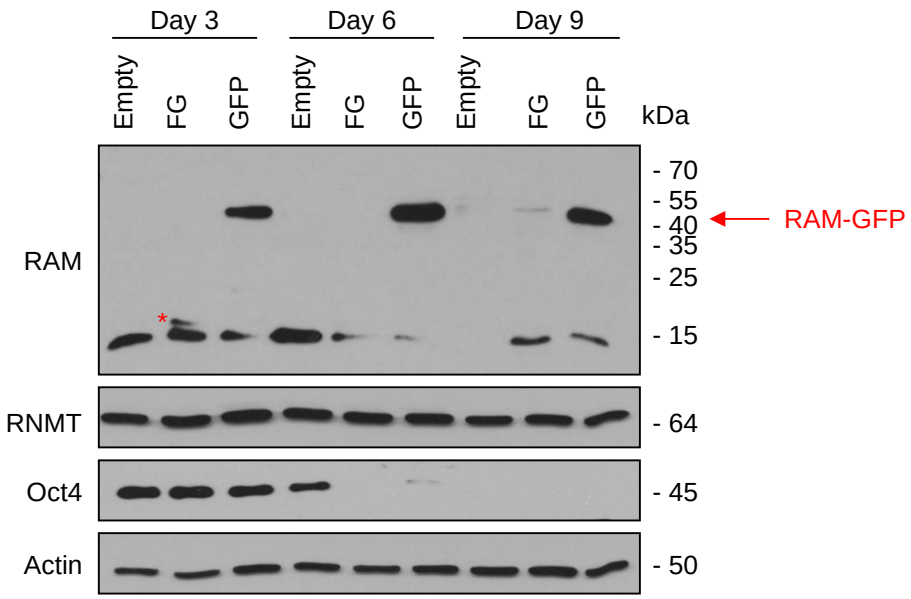


Figure 5.8: Expression of RAM-GFP is not reduced through neural differentiation.
In vitro neural differentiation protocol was performed with mESCs transfected with pPyPCAGGS-FgRAM (FG), RAMGFP (GFP) or with empty vector. Samples were collected on day 3, 6 and 9. Cell extracts were analysed by Western blot with the indicated antibodies. * Indicates the FG-RAM band. Actin serves as loading control.

Similarly to the control cells, the morphology of mESCs overexpressing RAM-GFP showed evidence of differentiation such as the production of protrusions that connected cells to each other (Figure 5.9a). Western blot analysis was employed to investigate whether and how overexpression of RAM could affect pluripotency factors expression during lineage commitment (Figure 5.9b). As expected, the RAM-GFP expression was maintained over the differentiation process. The upper band in the GFP blot indicated RAM-GFP whereas the lower one was representative of Sox1:GFP. The loss of Klf4 signal indicated that cells were losing the pluripotent phenotype towards the lineage commitment in agreement with the increase of Sox1:GFP signal. However, both Oct4 and Sox2 expression were maintained over time and did not decrease as observed in the control cell line. The experiment was repeated four times and quantification of Oct4 and Sox2 blots in control cells or cells over expressing RAM-GFP during the neural differentiation are shown (Figure 5.9c). Maintenance of RAM coincides with the maintenance of Oc4 and Sox2, which remained expressed until the last day of *in vitro* neural differentiation. This result is consistent with the siRNA study as it shows how RAM expression directly impacts Oct4 and Sox2 expression levels.

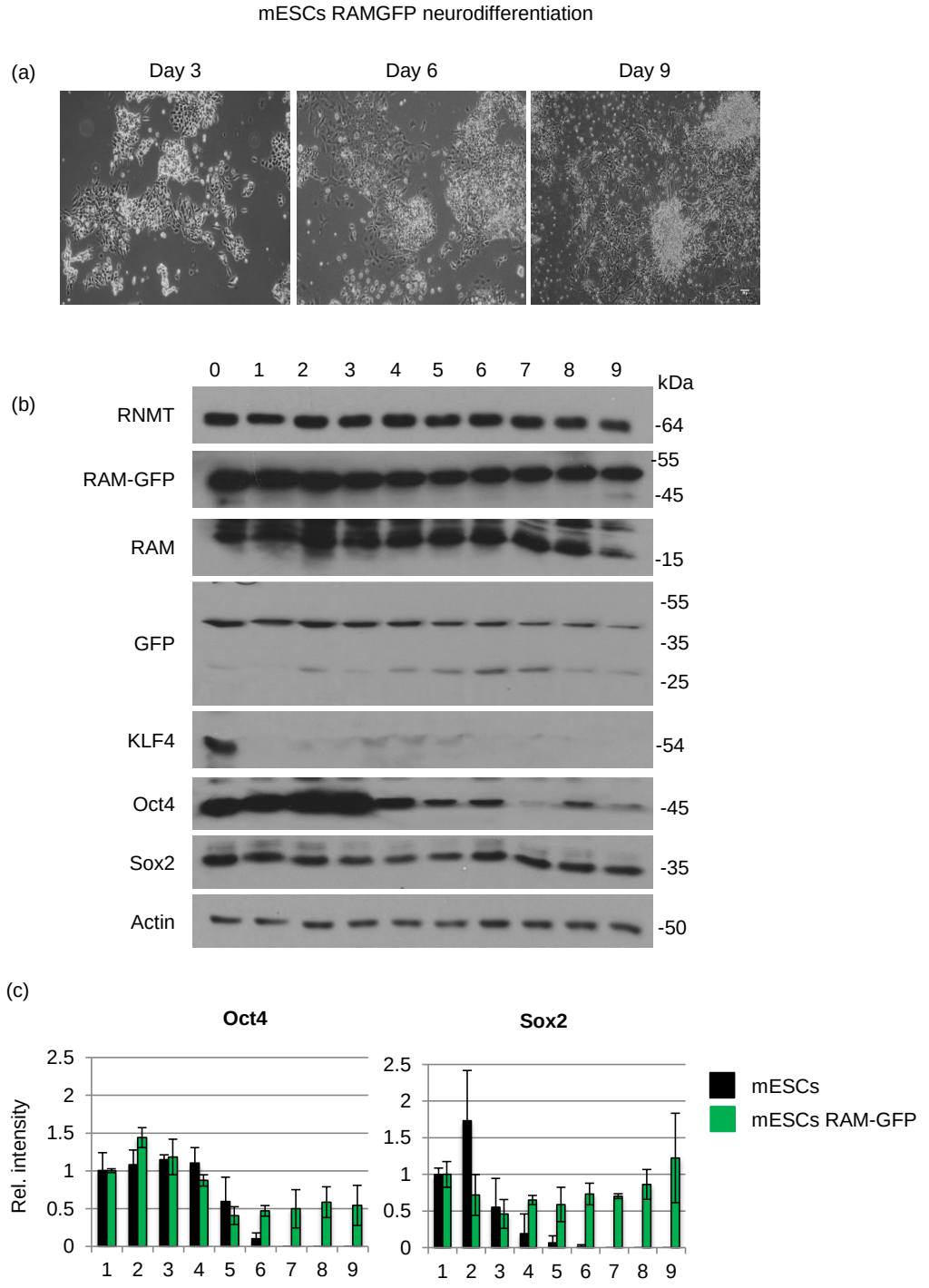


Figure 5.9: RAM overexpression maintains Oct4 and Sox2 expression during *in vitro* neural differentiation.

mESCs exogenously expressing RAM-GFP were seeded into a 10 cm dish coated with 0.1% gelatin in N2B27. Media was changed every second day. (a) Bright field of mESCs overexpressing RAM-GFP. Images were captured on days 3,6 and 9. Scale bar measures 20 µm. (b) Cell extracts collected at 24 hr intervals and analysed by Western blot with antibodies against the indicated proteins. Actin serves as loading control. (c) Quantification of western blots of Oct4 and Sox2 expression (relative to Actin expression) during the neural differentiation of mESCs and mESCs expressing exogenous RAM. Bar charts depicts the average values and relative standard deviation of four biological replicates. Quantification of western blot analyses was performed using Image J software.

The concomitant expression of neural marker, Sox:1-GFP and the maintenance of the pluripotency factors Oct4 and Sox2, prompted further investigation into the phenotype of mESCs over expressing RAM-GFP during the neural differentiation. Immunofluorescence was performed (Figure 5.10) and the expression of Oct4 and Tubulin-βIII, another neural marker, were investigated during the differentiation protocol in mESC control cells, used as positive control, and in mESCs over expressing RAM-GFP. As can be observed in the images, when mESCs are committed to the neural lineage, expression of Tubulin-βIII protein was observed within neurites. As previously described, Oct4 signal instead followed the opposite pattern, whereby high level of the transcription factors could be appreciated on day 0 and 3 of differentiation then decrease by day 6 and becoming further lower and similar to noise by day 9. On the other hand, in mESCs RAM-GFP, the Tubulin-βIII remained unspecific and never reached the high level of expression observed in the mESCs. Differently from the mESCs, was also the Oct4 expression that remained elevated until day 6 and on day 9 single cells were still expressing the protein. The microscopy data were consistent with the western blot analysis whereby Oct4 levels in mESCs over expressing RAM-GFP were maintained rather than decrease.

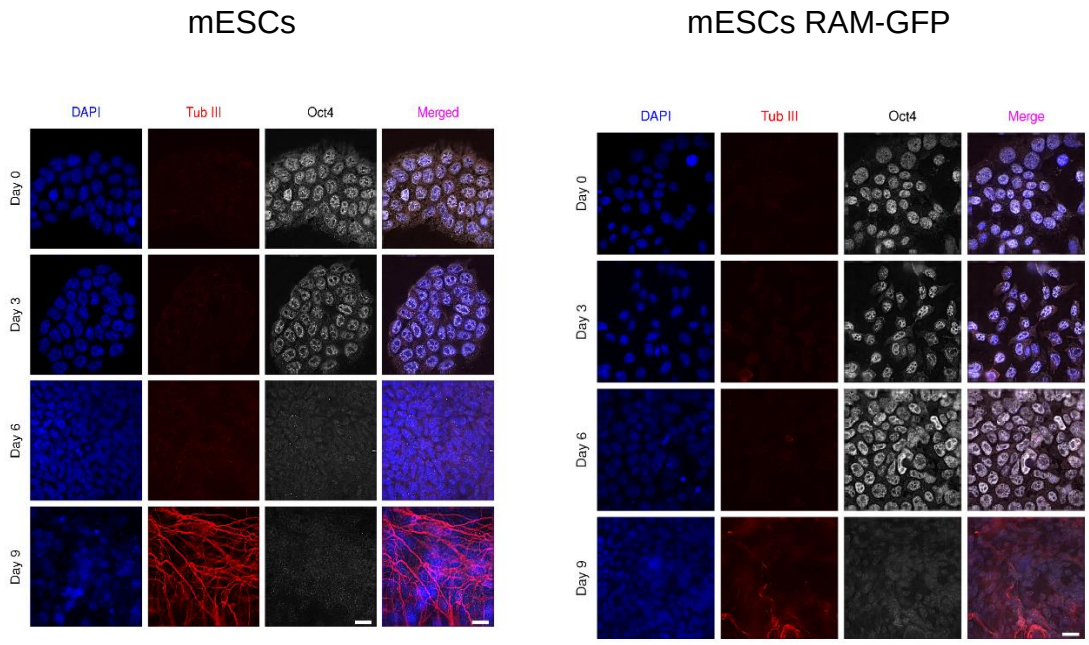


Figure 5.10: Immunofluorescence of mESCs and mESCs over expressing RAM-GFP shows that RAM expression delays the aquitance of neural phenotype compared to mESCs. Immunofluorescence microscopy was used to investigate whether cells that continue to express Oct4 are also co-expressing neural marker such as Tubulin-βIII. mESCs and mESCs over expressing RAM0GFP were seeded onto coverslip coated with Laminin in N2B27 media. Cells were fixed with 4% PFA and stained with the indicated antibodies on days 0, 3, 6 and 9. DAPI staining was used to detect nuclei. High-resolution images were collected with an imaging system (DeltaVision Restoration; Applied Precision) using a 60X/1.514 oil (Olympus) objective lens. Images were then processed using OMERO software. Scale bar measures 10µm. The above images are representative of two biological replicates (five fields were taken for each image).

Lastly, to see whether Oct4 and Sox2 expression was regulated at transcriptional level during the *in vitro* neural differentiation in the RAM-GFP cells, RT-qPCR analysis was carried out (Figure 5.11). RNMT and RAM transcript levels were maintained stable, whereas Oct4 mRNA levels, as described for the plain mESCs (Figure 4.4) were found to be gradually decreasing. On the contrary, Sox2 transcript levels remained steadily expressed over time, without neither halving as observed in the plain mESCs (Figure 4.4). The neuronal marker Nestin increased over time further confirming that mESCs despite maintaining RAM expression, gradually acquired a neural phenotype.

Thus these data showed that when RAM is overexpressed, despite the Oct4 protein expression was sustained during the protocol, the Oct4 mRNA continued to going down whereas Sox2 mRNA was consistent with the protein levels and remained steadily expressed. Considering the previous data on methyl capped mRNAs, it will be interesting to see whether the levels of methyl capped Oct4 and Sox2 transcripts in RAM-GFP neural differentiated are higher compared to control cells. This result confirms that RAM effect has a major impact on the translation of certain genes rather than on their transcription. All together these data describe a scenario whereby RAM directly affects the expression of Oct4 and Sox2 via regulating their methyl cap levels and thus their translational rate. Therefore these findings give an insight into the mechanism by which RAM expression contributes to maintain pluripotency in mESCs.

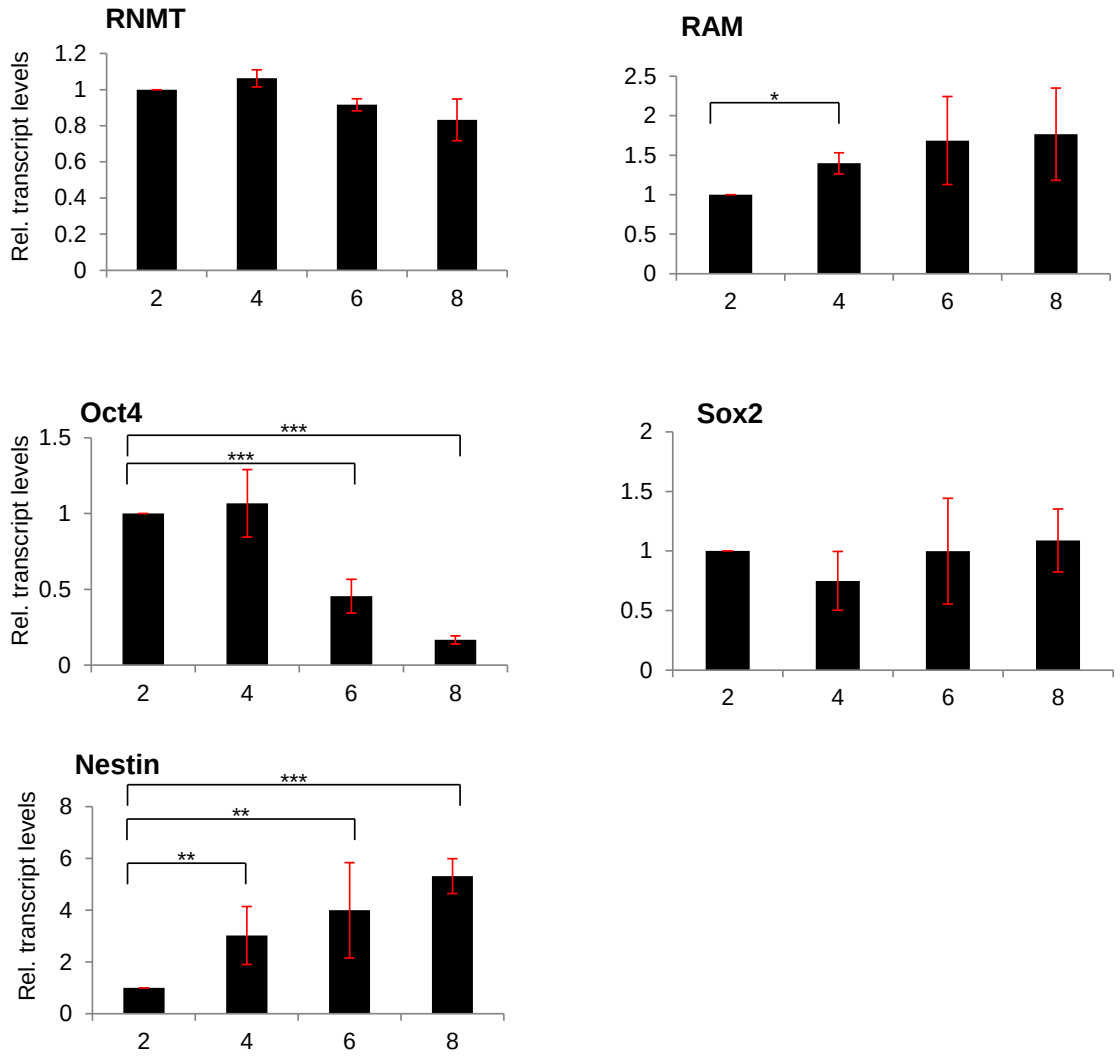


Figure 5.11: RAM overexpression affects protein levels of Oct4 and Sox2 but not transcripts.
mESCs expressing RAM-GFP were seeded into a 10 cm dish coated with 0.1% gelatin in N2B27 media for 9 days. Media was changed every second day. RNA extracted on day 2, 4, 6 and 8 of differentiation was analysed by RT-qPCR. The data were normalised to Actin mRNA levels. Bar charts depicts the average values and relative standard deviation of three biological replicates. Statistical significance ($p<0.05$) for all measurements was determined by two-tailed student's test, assuming unequal variance (n=3). * indicates $p<0.05$, ** indicates $p<0.005$.

5.3 Discussion

This chapter aimed to elucidate the function of RAM in mESCs and with this purpose loss and gain of function studies were performed. First, a preliminary Western blot analysis following RNMT or RAM knockdown showed that RAM depletion caused an impairment in Oct4 expression. Oct4 is a master pluripotency factor, required in critical amount as up or downregulation of Oct4 protein levels induce alternative lineage commitment (Niwa et al., 2000) . Due to the RAM effect on Oct4 expression and the possible consequences on loss of pluripotency, the overall effect of RAM on cell pluripotency was assessed. RAM knockdown was found to impair the ability of cells to form undifferentiated colonies and these findings combined together strongly support the hypothesis that high RAM expression in ESCs could have a functional role in the maintenance of pluripotency. This observation was further validated by the Western blot analysis of four of the main pluripotency markers. Results between independent biological replicates showed that RAM depletion particularly impairs Oct4, Sox2 and Klf4 and with lesser extent also Nanog protein levels. Surprisingly, RAM effect was not transcriptionally mediated, since none of the transcripts expressing Oct4, Sox2, Nanog and Klf4 were dramatically affected by its depletion. However, it is important to remember that RNMT possess the domain to catalyse the methylation of the cap but RAM harbours the RNA binding domain (Gonatopoulos-Pournatzis et al., 2011). Thus, the explanation for this could be that RAM depletion could affect the recruitment of mRNAs expressing the pluripotency markers to RNMT. This seems to be the case as observed by the methyl cap immunoprecipitations performed by Dr Victoria Cowling, which showed that differently from Actin, methyl cap levels of Oct4,

Sox2 and Klf4 transcripts were impaired after RAM depletion. These data correlate with the results obtained with polysome profiling where RAM knockdown was found to half polysome recruitment by mRNA transcripts expressing the pluripotency factors, thus explaining the lower protein expression. To discard the hypothesis that the RAM depletion causes a total reduction in RNA levels and consequently the decrease in polysome engagement, total RNA in fractions was calculated in cells treated with non targeting siRNA and with siRNA against RAM. The analysis was performed by Olga Suska and showed that RAM depletion led to a ~20% reduction in total RNA compared to control treated cells. Moreover, previous experiments showed that the overall yield of total RNA extracted from cells treated with non targeting control or siRNA against RAM was very similar (data not shown). Therefore, we can assume that the changes in translational rate of the pluripotency factors are due to RAM effect on their mRNA and not a lower RNA content.

The variability of the Nanog levels following RAM depletion observed by western blot may be explained by the following observation. It has been described that Nanog expression is highly heterogeneous within a stem cell population (Chambers et al., 2007), which may justify the discordant results obtained between the different experiments, and also in cells treated with non-targeting control, which showed low levels of Nanog expression. Furthermore, compensatory mechanisms could justify the evidence that Nanog expression has been not consistently affected by treatment with siRNA against RAM. Alternatively, RAM may have a preference towards Oct4, Sox2 and Klf4

transcripts but more studies are needed to sustain this latter hypothesis. On the other side, Oct4 functions as heterodimer with Sox2 in ESCs (Ambrosetti et al., 2000; Avilion et al., 2003; Masui et al., 2007), and additionally they regulate each other in an autoregulatory mechanism (Boyer et al., 2005; Chew et al., 2005). Thus, it does not surprise that RAM equally impacts on both transcription factors. Therefore, these data support a scenario where RAM specifically targets the methyl cap machinery to mRNAs for pluripotency markers to maintain their requisite levels of expression.

Considering the different effect of single siRNA against RAM, e.g. siRNA number 2 seemed the more effective, in the future it will be interesting to evaluate the efficiency of transfection of siRNAs. This can be addressed by quantifying the number of cells that are effectively transfected with the oligonucleotide sequence of the siRNA. Previous experiments in the lab were performed with rhodamine labelled siRNA against RNMT and RAM in human cells and showed that the siRNA uptake was 95% (unpublished data/data not shown). These validation need to be performed in mESCs.

Moreover, data shown in Figure 5.1 revealed that RNMT and RAM depletion cause and impairment on the other interacting partner although the reduction is less dramatic than that observed in human cell lines previously analysed (Gonatopoulos-Pournatzis et al., 2011). It will be interesting in the future to evaluate whether mouse RNMT and RAM have the same half-lives of the human orthologues (12 hr, unpublished data). In order to determine protein half-life, we can either treat cells with cycloheximide to block protein synthesis and

then detect the rate of loss of RNMT or RAM protein by western blot, or we can incubate methionine-starved cells with 35S methionine to label RNMT and RAM and then measure decay by SDS-PAGE and phosphor imaging. Unfortunately cycloheximide and methionine starvation treatments are toxic for mESCs and could not be used for the time length required to measure RNMT or RAM decay.

To confirm the effect of RAM on Oct4 and Sox2, mESCs expressing RAM-GFP were neural differentiated and showed that rather than decreasing over time as expected, Oct4 and Sox2 were stably maintained. Also in this context, the influence of RAM particularly on Oct4 was not transcriptional since its mRNA levels were gradually decreased. So it can be hypothesised that despite mRNA of Oct4 being downregulated, the pool of mRNA remaining was ensured to be translated.

The immunofluorescence between mESCs and mESCs over expressing RAM-GFP data clearly showed as contrary to mESCs, the Tubulin-βIII neural marker expression is delayed and minimal in the cells exogenously expressing RAM. It was worth to investigate whether those cells could be double positive and could co-express Oct4 and Tubulin-βIII but this was not the case as Tubulin-βIII signal was definitely lower and unspecific compared to mESCs. Therefore, overall, it may be speculated that over expression of RAM sustains the expression of Oct4 and Sox2, therefore resulting in a delay of the differentiation process. However, once the lineage commitment has been triggered by the external stimuli (media composition) RAM may not be enough to reverse the differentiation process as suggested by the neural phenotype observed by the

bright field, which resembled that one of the mESCs neural differentiated cell morphology, but also by the increase in Nestin mRNA levels and in Sox:1-GFP protein level. In the future it will be interesting to perform the neural differentiation for longer period of time and closely observed the delay described.

Moreover, it was not clear whether RAM upregulates one of the factors first, which then leads to higher expression of the other. Overall these data strongly support the previous results obtained with the siRNA study.

The neural differentiation protocol has been extensively repeated and it has been observed that the day RAM goes down may vary from one experiment to the other. Variations may be likely due to the different passage cells were used for the protocol. However, the dynamic of RAM regulation was always confirmed indicating that towards the last days of the neural differentiation the protein is downregulated. Additionally, sometimes upon expression of RAM-GFP, the expression of the endogenous RAM protein was markedly reduced in comparison to the exogenous one (Figure 5.8). This phenomenon has been previously observed and may indicate a negative feedback loop mechanism that regulates the physiological levels of RAM protein. More experimentation is required to validate and understand this process.

Thus data presented here support the hypothesis that RAM could preferentially recognise specific transcripts involved in the maintenance of pluripotency and recruit them to RNMT to allow the efficient methylation of their inverted mRNA guanosine cap. This assumption justifies the high expression levels of RAM in

stem cells, and as soon as cells differentiate and the global transcription decreases in favour of lineage specific markers genes, high levels of RAM are no longer required to sift through the transcriptional noise and thus is downregulated.

This chapter has described that RAM is the critical and fine tuning component within the methyltransferase complex. These data encourage further investigation of the global effects of alterations in RAM expression in mESCs. To address this a proteomics analysis following RAM depletion is currently being performed by Olga Suska in our laboratory.

6 Possible mechanism for regulation of RAM

6.1 Introduction

After RNA has been translated, the polypeptide chain undergoes folding and is often post-translationally modification to generate a final functionally active product. Post-translation modification by the covalent addition of chemical groups represents an important strategy through which cells diversify and extend protein function beyond what is dictated by gene transcripts (Uy & Wold, 1977). To further amplify this functional complexity, a plethora of PTMs including phosphorylation, ubiquitination, sumoylation, neddylation, acetylation, methylation, glycosylation and many more, have been reported. PTMs can be extremely dynamic and reversible. They can regulate the protein function affecting its activity, protein-protein interaction, localisation and stability (Seo & Lee, 2004). Recent large scale analyses have identified how key proteins in the pluripotency network are modified, especially phosphorylated, adding a new layer of regulation to their expression and activity (Wang *et al.*, 2014c; Cai *et al.*, 2012).

The previous chapters have described RAM as a critical component of the methyltransferase complex as it is involved in the regulation of pluripotency markers and is down-regulated during *in vitro* neural differentiation. So far, no evidence has been collected regarding the potential mechanism for regulation of RAM expression. Therefore, in order to have a complete picture of the biological importance of RAM in mESCs, the mechanism that regulates its expression was investigated. This chapter will be focused in exploring the possibility that RAM may be regulated via PTMs.

6.2 Results

6.2.1 RAM is degraded via proteasome degradation

In figure 4.4, it was shown that during *in vitro* neural differentiation, RAM transcripts, differently from its protein levels, are not down-regulated. A plausible hypothesis to justify the discrepancy between transcript and protein levels is that the *in vitro* neural differentiation process impacts the stability of RAM protein. In order to address this, the MG132 proteosomal inhibitor was used. Briefly, cells were seeded into N2B27 media for *in vitro* neural differentiation and were treated with 1 µM MG132 or DMSO for 3 hr prior to collection (Figure 6.1). On day 7, untreated cells or DMSO-treated showed a substantial reduction in RAM expression compared to day 5, which is consistent with what has been previously observed. On the contrary, the treatment with MG132 stabilised RAM as seen by the presence of a clear band. As expected, no effect was observed on the RNMT or Oct4 signals, the latter of which was reduced compared to day 5 confirming that cells underwent neural differentiation. This evidence established that RAM expression is reduced due to proteosomal degradation.

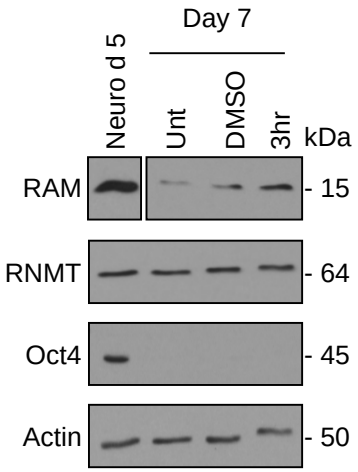


Figure 6.1: RAM loss during *in vitro* neural differentiation occurs via proteasomal degradation.
mESCs were cultured in a 10 cm dish coated with 0.1% gelatin in N2B27 media for 7 days. Cells were treated with 1µM MG132 or DMSO for 3 hr prior collection. Western blot analysis with antibodies against the indicated proteins. Samples of the RAM blot derive from the same gel. Actin serves as loading control. The above data are representative of two biological replicates.

6.2.2 RAM S36 is involved in degradation of RAM

Previous data have indicated that RAM is degraded during *in vitro* neural differentiation. It is well established that phosphorylation can often regulate the ubiquitin-mediated proteosomal degradation of target proteins (Hunter, 2007). The Serine-36 residue of RAM protein has been previously found to be phosphorylated (unpublished data, Thomas Gonatopoulos-Pournatzis), and as depicted in Figure 6.2, the amino acid sequence of RAM is well conserved from mice to humans, and so is the Serine-36 residue. In order to address if the phosphorylation of RAM at S36 is involved in the process of RAM degradation, the levels of this modification following *in vitro* neural differentiation were investigated using a phospho-specific antibody. Cells were seeded into N2B27 differentiation medium and samples collected every day (Figure 6.3). Western blot analysis clearly revealed that concomitant to the expected decrease in RAM levels, RAM phospho-S36 signal increased and remained highly expressed until day 9. The RNMT levels did not change and Oct4 expression was diminished as expected. Combining these observations so far, it can be hypothesised that during *in vitro* neural differentiation RAM is phosphorylated and then degraded.

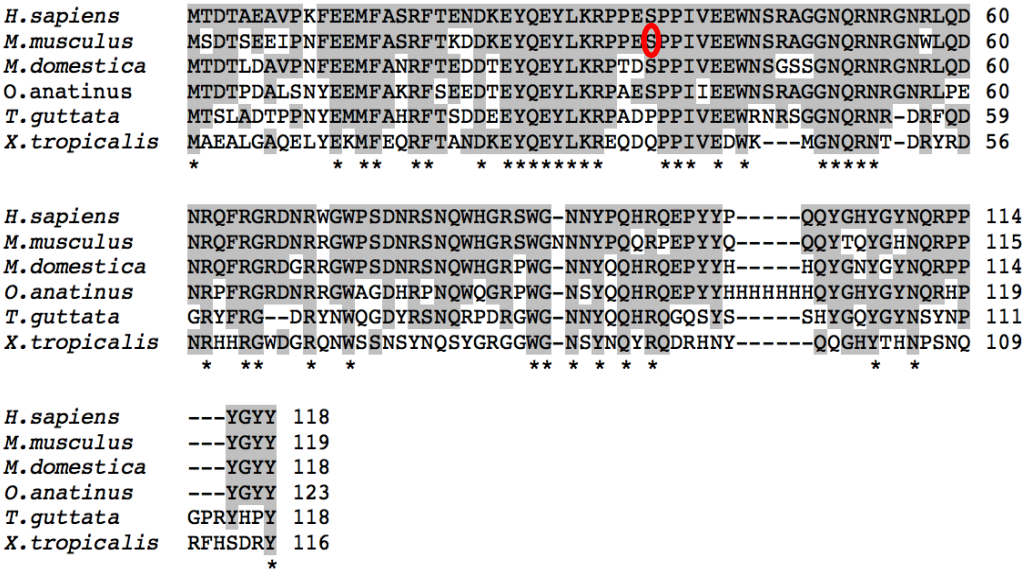


Figure 6.2: Amino acid sequence alignment of the *Homo sapiens* RAM protein and homologs in *Mus musculus*, *Monodelphis domestica, Ornithorhynchus anatinus, Tangara guttata, and Xenopus tropicalis*. ClustalW2 Multiple Alignment software was used for the alignment, selecting the default parameters. Amino acids identical in *H. sapiens* RAM protein and at least another species are highlighted in grey, and those identical in all species investigated are indicated (*).

As seen previously, RAM is expressed at low levels in MEFs compared to mESCs and mouse iPSCs (Figure 3.9). In order to investigate whether the phosphorylation of RAM coincides with its reduced expression also in MEFs, Western blot analysis was performed in mESCs, MEFs and mouse iPSCs (Figure 6.3b). In agreement with the previous data, the RAM phospho-S36 levels were higher in MEFs compared to mESCs and in mouse iPSCs, whereby the RAM phospho-S36 was undetectable. This evidence further supports the hypothesis that low levels of RAM are associated with its degradation via the phosphorylation at residue Serine-36.

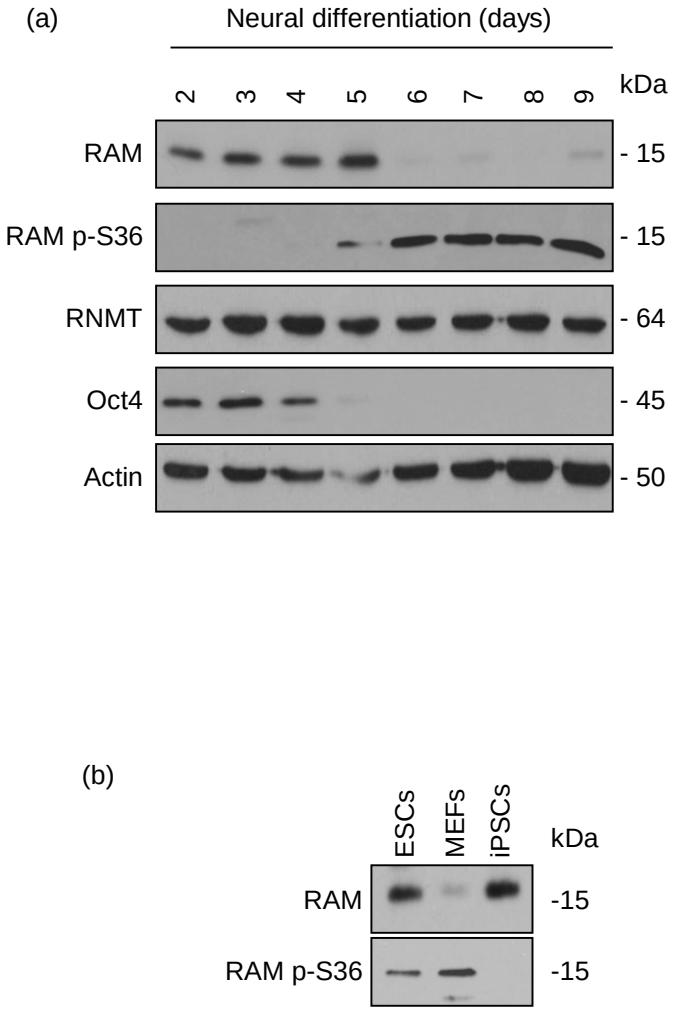


Figure 6.3: RAM is phosphorylated on Serine-36 during *in vitro* neural differentiation and in MEFs.
(a) Cells were seeded into a 10 cm dish pre-coated with 0.1% gelatin in N2B27 media for 9 days. Media was changed every second day, samples were collected at 24 hr intervals and analysed by Western blot with the antibodies against the indicated proteins. Actin serves as loading control. (b) Cell extracts from mESCs and MEFs and mouse iPSCs were analysed by Western blot to detect RAM phospho-S36. Actin serves as loading control. The above data are representative of three biological replicates.

6.2.3 RAM is regulated by Oct4.

In addition to this, the hypothesis that pluripotency factors could be involved in the regulation of RAM was also explored. Oct4 and Nanog expression was knocked down in mESCs by siRNA transfection (Figure 6.4). Briefly, cells were treated with pool of three siRNAs against Oct4, Nanog or a non-targeting control siRNA for 72 hr, following which samples were analysed by Western blot. Excitingly, treatment with Oct4 siRNA caused impairment in RAM expression levels whereas treatment with Nanog siRNA did not. It must be noted that impairment in Oct4 levels also led to a decrease in Nanog expression, thus the reduction in RAM levels could be a consequence of the depletion in levels of either of these two crucial pluripotency factors. However, Oct4 seems the more likely candidate. Sox2 and Klf4 also need to be tested in order to fully evaluate which member of the core pluripotency network influences the expression of RAM protein.

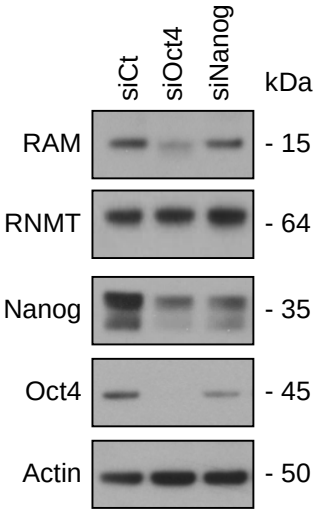


Figure 6.4: RAM expression is regulated by Oct4.
Western blot analysis performed on extracts from mESCs transfected with siRNA pools against Oct4, Nanog or non-targeting siRNA (siCt) for 72 hr with the antibodies against indicate proteins. Actin serves as loading control.

6.3 Discussion

This chapter aimed to elucidate the mechanisms regulating the expression of RAM in mESCs. Thomas Gonatopoulos-Pournatzis, former PhD student in our laboratory, found that in human cell lines, RNMT and RAM stabilise each other from proteosomal degradation. Contrary to this, the data presented here have shown that in mESCs undergoing *in vitro* neural differentiation, RAM is degraded by proteosomal activity while RNMT is not. These observations highlight how in stem cells, differently from the other cell lines previously analysed, RNMT and RAM are independently and differently regulated.

RAM was identified by mass spectrometry and the phosphorylation of Serine-36 was the only post translation modification detected. Other phosphoproteomics studies have further validated the phosphorylation at this residue (Goswami *et al.*, 2012; Wilson-Grady *et al.*, 2013). A specific RAM phosho-S36 antibody was raised by the DSTT (Division of Signal Transduction Therapy) service at the University of Dundee, the specificity of which was confirmed by Thomas Gonatopoulos-Pournatzis. Immunoprecipitation and orthophosphate labelling studies were carried out with cells expressing GFP, RAM-GFP and RAM-GFP S36A, where the latter has the Serine-36 residue mutated to Alanine and thus cannot longer be phosphorylated. Results showed that the RAM phospho-S36 antibody recognised only the wild type RAM and not the phospho null-mutant. The orthophosphate labelling revealed that RAM wild type was phosphorylated, whereas RAM S36A was not (data not shown). All together these data confirmed the presence of this post-translational modification and validated the tool to further investigate its role.

Data reported in this chapter showed that cells expressing low level of total RAM, such as MEFs and *in vitro* differentiated neural cells, show high levels of S36 phosphorylation on the remaining RAM molecules. Considering that the RAM antibody was raised against the full length RAM, these data strongly suggest that while the expression of total RAM decreases, the level of RAM phosphorylated at residue Serine-36 increases. In fact, the total RAM levels can be rescued by treatment with the proteosomal inhibitor MG132. These findings indicate that RAM is degraded via proteosomal degradation and the process is very likely triggered by its phosphorylation at Serine-36. It has already been described that phosphorylation can prime proteins for ubiquitination (Hunter, 2007). In order to investigate whether RAM is ubiquitinated, two cell lines were made expressing FG-RAM and phospho-null FG-RAM S36A. Experiments performed by Olga Suska revealed that wild-type RAM is modified by the addition of ubiquitin chains, whereas RAM S36A was not (data not shown). Therefore, our data combined together strongly support the idea that the phosphorylation at residue S36 is required to mediate RAM degradation via the ubiquitin-proteasome system (UPS).

Experiments performed by Dr. Dhaval Varshney in the lab have indicated that the phospho-null mutant RAM S36A or the phospho-mimic mutants RAM S36D/E do not impair its ability to enhance the methyltransferase activity of RNMT *in vitro*. These data need to be confirmed *in vivo*, however the recently resolved crystal structure of the RNMT-RAM complex and extensive mutagenesis of RAM have suggested that alterations in single residues are unlikely to dramatically impair its function (unpublished data, Dhaval Varshney).

The final data presented shed the light on the possibility of Oct4 being involved in the regulation of RAM expression. It is well established that Oct4 expression within a certain range is critical for stem-cell renewal, with any increase or decrease triggering differentiation to endoderm/mesoderm or trophectoderm, respectively (Niwa et al., 2000). In line with this observation, within one day of treatment with Oct4 siRNA, cells started to look different as they assumed a different morphology, indicating initiation of differentiation (data not shown). This strongly supports the finding that RAM is down-regulated when differentiation is triggered either during *in vitro* neural differentiation or due to loss of Oct4. However, the precise involvement of Oct4 in regulation of RAM expression still needs to be clarified.

7 Final discussion and future work

7.1 Final discussion

7.1.1 RNMT and RAM in mESCs

The addition of the methylguanosine cap at the 5' end of pre-mRNA is a critical event for the eukaryotic mRNA life cycle and consequently for cell viability (Chu & Shatkin, 2008; Mao et al., 1995; Tsukamoto et al., 1997). The co-transcriptional addition of the methyl cap is achieved by sequential activity of two enzymes, the latter of which, RNMT and its coactivating subunit RAM, methylates the cap in position N7. Increasing evidence has pointed out that the methylation of the cap, despite being so crucial for mRNA metabolism, is not a constitutive process but similarly to other steps within the gene expression, is subjected to regulation (Cole & Cowling, 2009; Cowling & Cole, 2007; Jiao et al., 2013). A peculiar gene expression program allows stem cells to self-renew and thereby maintains their pluripotency (Smith, 2001). The core pluripotency network lies at the top of the transcriptional regulatory hierarchy that activates genes required to maintain pluripotency meanwhile repressing lineage-specific genes (Young, 2011). Considering how critical the methyl cap synthesis is for gene expression, we decided to investigate whether the synthesis of the methyl cap plays a role in the maintenance of pluripotency.

The data presented in this thesis show that RNMT and RAM are highly expressed in mouse embryonic stem cells compared to differentiated cells. In addition to this, the expression of these two proteins is differentially regulated during development and differentiation as observed by their differential stoichiometry in various mouse organs.

As previously shown for TFIID and the RBP SON, when a protein is highly expressed in ESCs, it may likely be involved in the maintenance of pluripotency (Lu et al., 2013; Pijnappel et al., 2013). The evidence presented here also finds this to be the case for RNMT and RAM. In fact, the generation of iPS from MEFs restored the two proteins to high levels found in ESCs, indicating that the high RNMT and RAM expression is a feature of pluripotent stem cells.

Gel filtration and immunoprecipitation analyses revealed that in mESCs not all RNMT may be bound to RAM and the methyltransferase activity assay showed that elevated levels of RNMT and RAM result in a higher enzymatic activity in mESCs compared to MEFs. Considering that the brain exhibited high RNMT to RAM ratio, the regulation mechanism of the two proteins within the neural differentiation process was investigated. Following *in vitro* neural differentiation, RNMT clearly remained steadily expressed whereas RAM protein, but not transcript, levels were found to be gradually reduced. When RAM was discovered, it was described as an obligate component for the activity of the mammalian cap methyltransferase, whereby it binds, stabilises and increases RNMT activity both *in vitro* and *in vivo* (Gonatopoulos-Pournatzis et al., 2011). However, the exact mechanism by which RAM functions and its regulation remains unresolved. Therefore, data collected so far strongly encourage the investigation of how the RNMT and RAM complex functions in stem cells and why its components are differentially regulated.

7.1.2 RAM contribution to pluripotency

In order to address the role of RNMT and RAM complex in stem cells, gain or loss of function studies were performed. These revealed that RAM expression is integral to the control of the highly interconnected transcriptional network required for ESC pluripotency. Specifically high expression of RAM is required to maintain pluripotency as alterations in RAM levels directly affect important pluripotency factors. The important role of RAM is due to its ability to bind RNA. In fact, the central domain of RAM is characteristic of proteins with RNA binding activity (Bayer et al., 2005) and consistently to this, in our laboratory, it was shown that RAM was able to bind RNA *in vitro* whereas RNMT, despite possessing the catalytic domain to methylate the cap, did not (Gonatopoulos-Pournatzis et al., 2011). Based on methylguanosine cap immunoprecipitation and polysome profiling following RAM depletion, whereby the first technique measures the amount of methyl capped mRNA of a given transcript and the second gives an indication of its relative translation, a substantial reduction in the methylated capped RNA and consequently protein levels of Oct4, Sox2 and Klf4 was observed. A minor degree of impairment was described Nanog. Parallel studies were performed with mESCs exogenously expressing RAM throughout neural differentiation. During the *in vitro* neural differentiation, concomitant to RAM being stably expressed, also Oct4 and Sox2 protein levels were unexpectedly maintained. These data together highlight RAM effect mainly on Oct4 and Sox2 protein expression.

This finding adds another piece to the complex puzzle of transcriptional and translational regulation in stem cells. It has been published that ESCs have a

low-levels of global transcription, expressing regions of the genome that are normally silenced in differentiated cells (Efroni et al., 2008). As previously described, the core pluripotency factors can be found along with PRC2 at bivalent domains (Bernstein et al., 2006; Lee et al., 2006). Although originally thought to be silent, recent findings suggest that some of these polycomb-repressed genes are in fact transcribed in ESCs (Brookes & Pombo, 2009). Nonetheless, the eventual fate of these transcripts is unclear because despite being capped (not clear whether methyl capped), they are not translated. It is possible that for some genetic loci, it may be energetically more convenient to recruit the whole transcriptional machinery to the transcription site and then regulate the fate of the transcripts (Pombo, personal communication). The data obtained in this study allow the speculation that transcripts required for pluripotency are recruited by RAM and presented to RNMT for the methylation of their cap, thus resulting in the final gene product being expressed. It can be hypothesised that transcripts from bivalent genes which should be repressed, are not recruited by RAM to RNMT, thus their cap structure is not methylated, and consequently targeted by decapping enzymes. This model will agree with the observed presence of decapping machineries that specifically degrade unmethylated capped transcripts (Jiao et al., 2013; Song et al., 2013). Moreover, the finding that undifferentiated embryonic cell transcription factor 1 (Utf1) recruits DCp1, a component of the mRNA decapping machinery, to bivalent leaky promoters mediating the cytoplasmic degradation of their relative mRNAs (Jia et al., 2012), further supports this model.

The reduced transcriptional noise in differentiated cells (Efroni et al., 2008; Meshorer & Misteli, 2006) reduces the requirement for increased specificity of cap methylation, thus diminishing the requirement for elevated RAM expression. In order to confirm this hypothesis, m7G immunoprecipitation and polysome profiling need to be extended to other transcription factors of the pluripotency network as well as lineage specific genes. For what concerns translation in stem cells, two studies based on polysome and ribosome profiling revealed that stem cells are characterized by low translation levels with surfeit of free ribosomes. On the contrary, differentiation in embryoid bodies induces an anabolic switch that leads to increase their Golgi apparatus and rough endoplasmic reticulum coordinated with increase in the ratio Cytoplasm/Nucleus ratio and synthesis of ribosomal proteins (Ingolia *et al.*, 2011; Sampath et al., 2008). Consequently, EBs present higher engagement of polysome loaded onto mRNAs and thus protein synthesis. In line with our model, these evidences together sustain that in stem cells high levels of transcriptional noise are coupled to a parsimonious translation whereby the fate of specific transcripts is decided by targeted cap methylation.

7.1.3 RAM-Oct4-Sox2 network

Besides this, the findings that Oct4 and Sox2 are specifically sensitive to RAM expression levels are further supported by other studies where RNA binding proteins display preferential activity towards certain transcripts. For example, the Thoc5 subunit of the THO complex specifically mediates the export and expression of Nanog and Klf4, but not Oct4 (Wang et al., 2013).

Considering that Oct4 is a central pluripotency marker and its expression levels need to be within a certain threshold concentration (Niwa et al., 2000), it is plausible to think that different and several regulatory mechanisms collaborate to maintain its required expression levels. This study describes cap methylation as another important step through which the expression of pluripotency markers could be tightly and specifically regulated to maintain stem cells in a pluripotent state.

The previous data explains how alterations in RAM levels leads mainly to perturbation in Oct4 and Sox2 protein levels. Considering that during the *in vitro* neural differentiation RAM levels are dramatically reduced, and so the Oct4 and Sox2 levels, and when RAM is exogenously expressed, Oct4 and Sox2 expression is maintained throughout the differentiation process, it can be speculated that RAM reduction contributes to the differentiation process. On the other side, Oct4 depletion was shown to impair RAM expression, (current studies in the lab are showing that Sox2 has the same effect), therefore combining these evidence together seems that a RAM-Oct4-Sox2 are co-dependent. However, all together these data make difficult to distinguish whether RAM downregulation is cause or consequence of the differentiation process. In order to discern between the two possibilities future experiments will deplete RAM to than perform the *in vitro* neural differentiation protocol to see whether and how the dynamic of the differentiation process is affected (Figure 7.1).

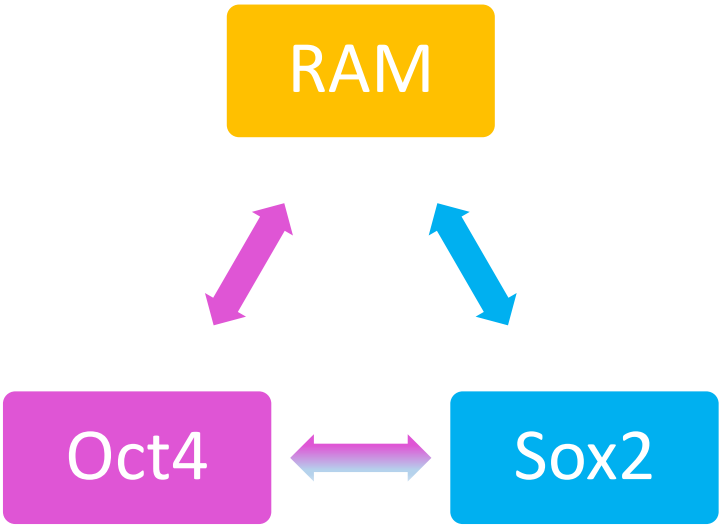


Figure 7.1: Is RAM lost cause or consequence of differentiation?
Specifically high expression of RAM is required to maintain pluripotency as alterations in RAM levels directly affect important pluripotency factors, particularly Oct4 and Sox2. Gain and loss of function studies revealed RAM and Oct4 and Sox2 ae co-dependent. More work is needed to elucidate the interplay between RAM and the pluripotency factors.

A recent study has shown that the abundance of SAM is required for maintenance of pluripotency in mESCs and its levels are regulated by threonine metabolism. Therefore, when authors depleted threonine from the culture or impair the threonine dehydrogenase in mESCs, it leads to decreased SAM levels that causes a reduction in H3K4me3, which ultimately induced differentiation (Shyh-Chang *et al.*, 2013). This observation has also been confirmed in hESCs and hiPSCs, whereby SAM was demonstrated to also be required for self-renewal (Shiraki *et al.*, 2014). The scenario that emerges is that SAM is required in abundance in stem cells but it is highly used and therefore limiting. In fact, both works pointed out that reduction of SAM levels leads to decrease in histone methylation, however, in light of our findings, we can speculate that since SAM is limiting and highly competed for, high RAM levels will enable RNMT to recruit SAM more efficiently and consequently enable optimal cap methylation.

Data presented here corroborate the increasing number of evidence that depicts the methylation of the cap as a more regulated modification than was originally thought (Cowling & Cole, 2010; Jiao et al., 2010). An important aspect of the research presented here is that part of the results obtained with mESCs have also been validated in hESCs. In fact, human embryonic stem cells and human iPS also exhibit high levels of RNMT and RAM, the latter of which is similarly lost during *in vitro* neural differentiation of hESCs. Therefore, our findings in mESCs represent a promising model for a better understanding of the regulation of pluripotency in hESCs, and can also be relevant for the generation of hiPSCs. The data previously discussed and presented here

emphasise the critical role of RNA and consequently RNA binding proteins in the maintenance of pluripotency.

7.1.4 Regulation of RAM

Data presented in chapter 6 has paved the way to further investigate the mechanism of regulation of RAM expression. As previously discussed, when stem cells differentiate, they down regulate RAM expression via proteasomal degradation targeted through phosphorylation at Serine-36 residue. Among the different degradation pathways, the 26 proteosome, which degrades ubiquitinated proteins to small peptides, is one of the most characterized. Ubiquitination consist in the covalent attachment of ubiquitin (Ub) to a substrate through an enzymatic cascade involving ubiquitin-activating (E1), ubiquitin-conjugating (E2), and ubiquitin-ligating (E3) enzymes. Briefly the E3 ligase binds to both substrate and E2 to facilitate the ubiquitin transfer from E2 to the substrate (Deshaies & Joazeiro, 2009). Evidence collected over the last decades demonstrated a crosstalk between phosphorylation and ubiquitination (Hunter, 2007). In fact, phosphorylation can regulate ubiquitination in three main mechanisms: firstly directing regulating the activity of the E3 ligase responsible for Ub transfer. Secondly it can promote recognition by an E3 ligase generating a phosphodegron and thirdly, phosphorylation can regulate the subcellular localization of substrate/ligase interaction (Hunter, 2007). The hypothesis that RAM is ubiquitinated is currently being tested in the lab.

Moreover, preliminary experiments have suggested that RAM can be phosphorylated by ERK *in vitro* (unpublished data, Lize Wasmus). ERK is the final effector of the MAPK pathway and although recent findings support the opposite (Hamilton & Brickman, 2014), multiple evidence have revealed that ERK is required to sustain neurodifferentiation (Kunath *et al.*, 2007; Stavridis *et al.*, 2007). Thus, it is plausible to hypothesise that RAM is phosphorylated by ERK, and that this aids the ERK-mediated neurodifferentiation program. Therefore, future experiments will verify this possibility.

7.2 Future work

This PhD project offers a new prospective into the regulation of pluripotency in stem cells. However, from this study many new questions arise, which need to be addressed in order to have a clear picture regarding the contribution of mRNA cap methylation in the maintenance of stem cell pluripotency.

1. **CLIP analysis combined with RNA sequencing**

   Our current hypothesis is that RAM functions as a sieve in stem cells to specifically recruit mRNAs expressing pluripotency factors to RNMT, which then methylates their cap. Therefore, it is important to understand how RAM is able to distinguish and recognise specific transcripts. Currently, Dr. Dhaval Varshney is optimizing a high throughput sequencing following cross-linking and immunoprecipitation (HITS-CLIP) protocol to identify specific sequences bound to the RNMT-RAM complex.

In the future, we will use this technique to analyse how the RNA binding profile for RNMT and RAM complex changes between mESCs and *in vitro* neural differentiated cells. This will enable a direct analysis of any specific targeting of cap methylation and its contribution to the pluripotency program. We would also like to evaluate the impact of the dramatic reduction in RAM expression levels on RNMT RNA-binding.

2. **Proteomic analysis following RAM depletion**

   In order to have a broader view on the RAM-dependent proteome in stem cells, Olga Suska is currently performing proteomic studies following depletion of RAM with siRNA.

3. **Establishing the mechanism that regulates RAM expression in stem cells**

   Data presented here shed the light on RAM being degraded via proteasomal degradation in a mechanism that requires the phosphorylation of RAM at Serine-36 residue. Further experiments will be focused on unrevealing how exactly the mechanism occurs. Moreover, preliminary studies have shown that Oct4 (maybe along with Nanog) impairs RAM expression. Consistently to the post-transcriptional regulation of RAM, published ChIP-seq datasets do not show Oct4 nor Nanog occupancy on the RAM promoter (Mathur *et al.*, 2008; Chen et al., 2008), therefore possibly Oct4 will activate a kinase that in turn will phosphorylate RAM and thus label it for degradation.

Nevertheless, E2F1 has clearly been shown to occupy RAM promoter, whereas Myc, Esrrb, Klf4, n-Myc and Zfx have been detected at the RNMT promoter (Chen et al., 2008). E2F1 and c-Myc are part of the c-Myc-centric module, which as previously mentioned, is typical of genes that are actively transcribed in stem cells (Chen et al., 2008). However, ChIP data showed that SIN3b, a cofactor of the RE1 silencing transcription factor (REST) occupies the RAM, but not the RNMT promoter (Yu *et al.*, 2011). Therefore RNMT, which contains the catalytic subunit to methylate the cap, is constantly active whereas RAM can be activated or repressed. Thus, the possibility that also a transcriptional mechanism could contribute to the expression levels of RAM further emphasises the role of RAM as a critical component of the methyltransferase complex in stem cells. Future studies are required to validate this hypothesis.

4. Exploring the role of RNMT and RAM in other organs

Song and colleagues observed that Dcp2 (a decapping enzyme) was differentially expressed in mouse organs and showed that knockout of Dcp2 had no adverse consequences in mice. Taking these results into account, they suggested that Dcp2 could not be the sole decapping protein. Indeed they have identified another decapping enzyme called Nud16, which possibly works on different subset of mRNAs (Song et al., 2010). Similarly, the differential stoichiometry of RNMT and RAM in different mouse organs supports the possibility that other interacting partners may be involved in the RNMT and RAM complex. This is further

support by our gel filtration and immuprecipitation data that clearly suggest that not all RNMT is bound to RAM and that the RNMT and RAM complex migrates higher than the expected molecular weight, indicating that other partners may take part in the complex. Another hypothesis is that once RAM is down regulated during *in vitro* neural differentiation, other RAM-like proteins may function in the opposite way and thus present lineage specific mRNAs to RNMT. Therefore, similarly to Song approach, RNMT and RAM KO mice will be generated to investigate whether other proteins take part in the methyltransferase complex. Moreover, based on the results obtained with different organs, it is likely that some organs will be more affected than others by RNMT and RAM KO. For example, the heart represents an extremely interesting case study since the RNMT and RAM pattern expression is exactly opposite to what has been described so far in neurons, whereby RNMT expression is reduced and RAM is present at elevated levels. Therefore, in the future, it will be also interesting to study the expression pattern of these proteins during mESCs differentiation into cardiomyocytes

Since as previously mentioned, is not easy to discern whether loss of RAM is cause or consequence of differentiation, the KO mice will help to clarify the importance of the methyltransferase complex during development. In fact, only two main works have been performed during early development. The first was carried out in sea urchin *Strongylocentrotus purpuratus* and showed that N7 ribose methylation occurs following fertilization but prior to the two cell stage (Caldwell & Emerson, 1985).

Later studies on *in vitro* RNA injected in *X. leavis*, showed that the activity of cytoplasmic guanine-7-methyltransferase increases during oocyte maturation and stimulates translation of unmethylated cap. This process is promoted by progesterone and the increase in methyltransferase activity is concurrent with nuclear breakdown (Gillian-Daniel et al., 1998).

The model proposed in this work, further support the relation between methyl cap formation and protein translation that seems to be specifically regulated by RAM in embryonic stem cells. eIF4E and c-Myc are fundamental for translation. The former is the cap-binding initiation factors, required for efficient and correct translation (Topisirovic et al., 2011). Whereas c-Myc is involved in numerous processes, amongst which it upregulates the methyl cap formation of its target genes (Cowling & Cole, 2010), which lead to cell proliferation. No data are available on eIF4E KO mice, possibly because they are not viable whereas Myc KO mice are lethal between day 9.5 or 10 (Davis *et al.*, 1993). Therefore, considering the intimately link between methyl cap moiety and translation, it will be extremely interesting to see how the KO RNMT and RAM mice behave. Possibly, more dramatic consequences could be observed in case of double KO of RNMT and RAM simultaneously.

Overall, this pioneering work has revealed the regulation of mRNA cap methylation as an additional mechanism that contributes to maintain stem cell pluripotency (Figure 7.2).

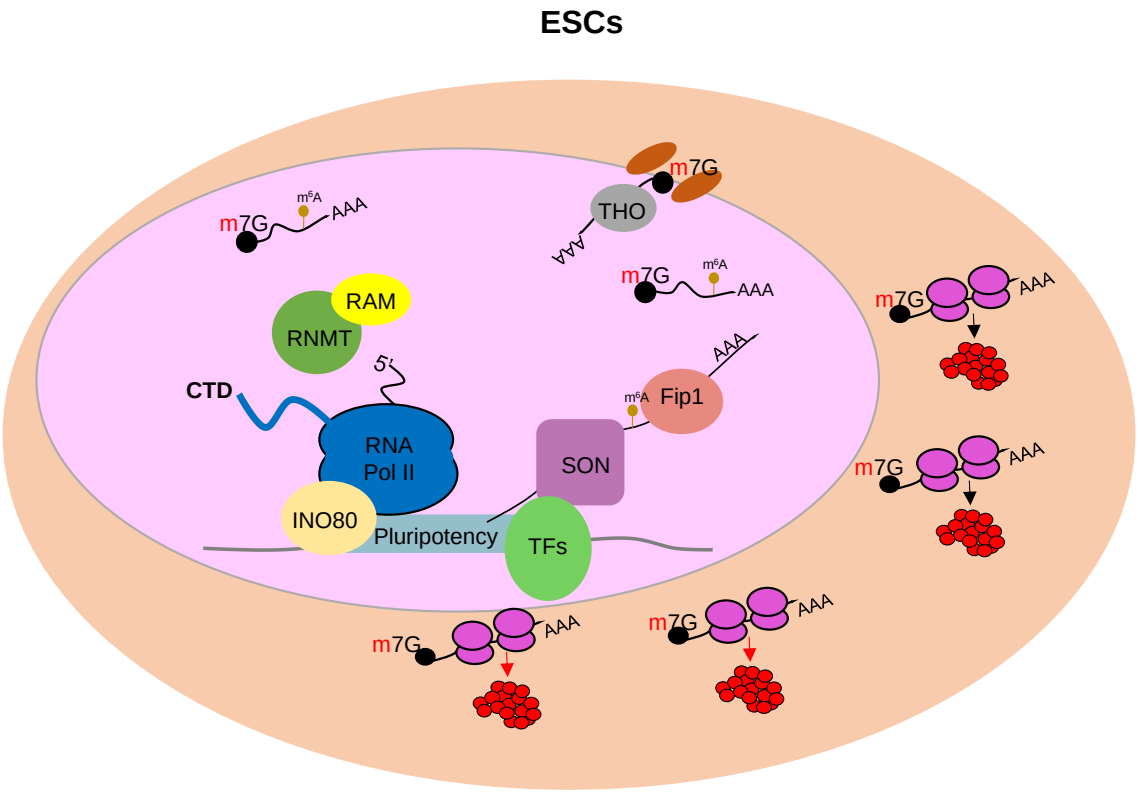


Figure 7.2: Cap methylation of mRNAs adds a new step of regulation of the ESCs state. Additionally to other mechanisms previously described, cap methylation joins the long list of event that stem cells use to maintain the undifferentiated state. RAM functioning as sieve, present the mRNA of pluripotency factors to RNMT that will proceed with the methylation of their cap and thus promoting their expression. Depletion of RAM results in reduction of protein levels of the main transcription factors.

8 Bibliography

ABRANCHES, E., BEKMAN, E. and HENRIQUE, D. (2013). Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. *PLoS One,* 8**,** e59928.

ABRANCHES, E., SILVA, M., PRADIER, L., SCHULZ, H., HUMMEL, O., HENRIQUE, D. and BEKMAN, E. (2009). Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. *PLoS One,* 4**,** e6286.

ALTMANN, M., EDERY, I., SONENBERG, N. and TRACHSEL, H. (1985). Purification and characterization of protein synthesis initiation factor eIF-4E from the yeast Saccharomyces cerevisiae. *Biochemistry,* 24**,** 6085-9.

ALTMANN, M., HANDSCHIN, C. and TRACHSEL, H. (1987). mRNA cap-binding protein: cloning of the gene encoding protein synthesis initiation factor eIF-4E from Saccharomyces cerevisiae. *Mol Cell Biol,* 7**,** 998-1003.

AMBROSETTI, D. C., BASILICO, C. and DAILEY, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol Cell Biol,* 17**,** 6321-9.

AMBROSETTI, D. C., SCHOLER, H. R., DAILEY, L. and BASILICO, C. (2000). Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J Biol Chem,* 275**,** 23387-97.

AREGGER, M. and COWLING, V. H. (2013). Human cap methyltransferase (RNMT) N-terminal non-catalytic domain mediates recruitment to transcription initiation sites. *The Biochemical journal,* 455**,** 67-73.

AVILION, A. A., NICOLIS, S. K., PEVNY, L. H., PEREZ, L., VIVIAN, N. and LOVELL-BADGE, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev,* 17**,** 126-40.

BAIN, G., KITCHENS, D., YAO, M., HUETTNER, J. E. and GOTTLIEB, D. I. (1995). Embryonic stem cells express neuronal properties in vitro. *Dev Biol,* 168**,** 342-57.

BALATSOS, N. A., NILSSON, P., MAZZA, C., CUSACK, S. and VIRTANEN, A. (2006). Inhibition of mRNA deadenylation by the nuclear cap binding complex (CBC). *J Biol Chem,* 281**,** 4517-22.

BATISTA, P. J., MOLINIE, B., WANG, J., QU, K., ZHANG, J., LI, L., BOULEY, D. M., LUJAN, E., HADDAD, B., DANESHVAR, K., CARTER, A. C., FLYNN, R. A., ZHOU, C., LIM, K. S., DEDON, P., WERNIG, M., MULLEN, A. C., XING, Y., GIALLOURAKIS, C. C. and CHANG, H. Y. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell,* 15**,** 707-19.

BAYER, T. S., BOOTH, L. N., KNUDSEN, S. M. and ELLINGTON, A. D. (2005). Arginine-rich motifs present multiple interfaces for specific binding by RNA. *RNA,* 11**,** 1848-57.

BELANGER, F., STEPINSKI, J., DARZYNKIEWICZ, E. and PELLETIER, J. (2010). Characterization of hMTr1, a human Cap1 2'-O-ribose methyltransferase. *J Biol Chem,* 285**,** 33037-44.

BEN-DAVID, U. and BENVENISTY, N. (2011). The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nat Rev Cancer,* 11**,** 268-77.

BENTLEY, D. L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr Opin Cell Biol,* 17**,** 251-6.

BERNSTEIN, B. E., MIKKELSEN, T. S., XIE, X., KAMAL, M., HUEBERT, D. J., CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S. L. and LANDER, E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell,* 125**,** 315-26.

BOTH, G. W., BANERJEE, A. K. and SHATKIN, A. J. (1975). Methylation-dependent translation of viral messenger RNAs in vitro. *Proc Natl Acad Sci U S A,* 72**,** 1189-93.

BOTQUIN, V., HESS, H., FUHRMANN, G., ANASTASSIADIS, C., GROSS, M. K., VRIEND, G. and SCHOLER, H. R. (1998). New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev,* 12**,** 2073-90.

BOYER, L. A., LEE, T. I., COLE, M. F., JOHNSTONE, S. E., LEVINE, S. S., ZUCKER, J. P., GUENTHER, M. G., KUMAR, R. M., MURRAY, H. L., JENNER, R. G., GIFFORD, D. K., MELTON, D. A., JAENISCH, R. and YOUNG, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell,* 122**,** 947-56.

BRAUNSCHWEIG, U., GUEROUSSOV, S., PLOCIK, A. M., GRAVELEY, B. R. and BLENCOWE, B. J. (2013). Dynamic integration of splicing within gene regulatory pathways. *Cell,* 152**,** 1252-69.

BROOKES, E. and POMBO, A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. *EMBO Rep,* 10**,** 1213-9.

CAI, N., LI, M., QU, J., LIU, G. H. and IZPISUA BELMONTE, J. C. (2012). Post-translational modulation of pluripotency. *J Mol Cell Biol,* 4**,** 262-5.

CALDWELL, D. C. and EMERSON, C. P., JR. (1985). The role of cap methylation in the translational activation of stored maternal histone mRNA in sea urchin embryos. *Cell,* 42**,** 691-700.

CALERO, G., WILSON, K. F., LY, T., RIOS-STEINER, J. L., CLARDY, J. C. and CERIONE, R. A. (2002). Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. *Nat Struct Biol,* 9**,** 912-7.

CAO, Q., PADMANABHAN, K. and RICHTER, J. D. (2010). Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition. *RNA,* 16**,** 221-7.

CARTWRIGHT, P., MCLEAN, C., SHEPPARD, A., RIVETT, D., JONES, K. and DALTON, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development,* 132**,** 885-96.

CHAMBERS, I., COLBY, D., ROBERTSON, M., NICHOLS, J., LEE, S., TWEEDIE, S. and SMITH, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell,* 113**,** 643-55.

CHAMBERS, I., SILVA, J., COLBY, D., NICHOLS, J., NIJMEIJER, B., ROBERTSON, M., VRANA, J., JONES, K., GROTEWOLD, L. and SMITH, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature,* 450**,** 1230-4.

CHANG, J. H., JIAO, X., CHIBA, K., OH, C., MARTIN, C. E., KILEDJIAN, M. and TONG, L. (2012). Dxo1 is a new type of eukaryotic enzyme with both

decapping and 5'-3' exoribonuclease activity. *Nature structural & molecular biology,* 19**,** 1011-7.

CHANGELA, A., HO, C. K., MARTINS, A., SHUMAN, S. and MONDRAGON, A. (2001). Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *EMBO J,* 20**,** 2575-86.

CHAU, V., TOBIAS, J. W., BACHMAIR, A., MARRIOTT, D., ECKER, D. J., GONDA, D. K. and VARSHAVSKY, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science,* 243**,** 1576-83.

CHEN, C. Y. and SHYU, A. B. (2011). Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip Rev RNA,* 2**,** 167-83.

CHEN, X., XU, H., YUAN, P., FANG, F., HUSS, M., VEGA, V. B., WONG, E., ORLOV, Y. L., ZHANG, W., JIANG, J., LOH, Y. H., YEO, H. C., YEO, Z. X., NARANG, V., GOVINDARAJAN, K. R., LEONG, B., SHAHAB, A., RUAN, Y., BOURQUE, G., SUNG, W. K., CLARKE, N. D., WEI, C. L. and NG, H. H. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell,* 133**,** 1106-17.

CHENG, H., DUFU, K., LEE, C. S., HSU, J. L., DIAS, A. and REED, R. (2006). Human mRNA export machinery recruited to the 5' end of mRNA. *Cell,* 127**,** 1389-400.

CHEW, J. L., LOH, Y. H., ZHANG, W., CHEN, X., TAM, W. L., YEAP, L. S., LI, P., ANG, Y. S., LIM, B., ROBSON, P. and NG, H. H. (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol,* 25**,** 6031-46.

CHIANG, P. K., GORDON, R. K., TAL, J., ZENG, G. C., DOCTOR, B. P., PARDHASARADHI, K. and MCCANN, P. P. (1996). S-Adenosylmethionine and methylation. *FASEB J,* 10**,** 471-80.

CHIN, M. H., MASON, M. J., XIE, W., VOLINIA, S., SINGER, M., PETERSON, C., AMBARTSUMYAN, G., AIMIUWU, O., RICHTER, L., ZHANG, J., KHVOROSTOV, I., OTT, V., GRUNSTEIN, M., LAVON, N., BENVENISTY, N., CROCE, C. M., CLARK, A. T., BAXTER, T., PYLE, A. D., TEITELL, M. A., PELEGRINI, M., PLATH, K. and LOWRY, W. E. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell,* 5**,** 111-23.

CHLEBOWSKI, A., LUBAS, M., JENSEN, T. H. and DZIEMBOWSKI, A. (2013). RNA decay machines: the exosome. *Biochim Biophys Acta,* 1829**,** 552-60.

CHO, E. J., TAKAGI, T., MOORE, C. R. and BURATOWSKI, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev,* 11**,** 3319-26.

CHU, C., DAS, K., TYMINSKI, J. R., BAUMAN, J. D., GUAN, R., QIU, W., MONTELIONE, G. T., ARNOLD, E. and SHATKIN, A. J. (2011). Structure of the guanylyltransferase domain of human mRNA capping enzyme. *Proc Natl Acad Sci U S A,* 108**,** 10104-8.

CHU, C. and SHATKIN, A. J. (2008). Apoptosis and autophagy induction in mammalian cells by small interfering RNA knockdown of mRNA capping enzymes. *Mol Cell Biol,* 28**,** 5829-36.

CHUANG, T. W., CHANG, W. L., LEE, K. M. and TARN, W. Y. (2013). The RNA-binding protein Y14 inhibits mRNA decapping and modulates processing body formation. *Mol Biol Cell,* 24**,** 1-13.

CLEVERS, H. (2005). Stem cells, asymmetric division and cancer. *Nat Genet,* 37**,** 1027-8.

COHEN, N., SHARMA, M., KENTSIS, A., PEREZ, J. M., STRUDWICK, S. and BORDEN, K. L. (2001). PML RING suppresses oncogenic transformation by reducing the affinity of eIF4E for mRNA. *EMBO J,* 20**,** 4547-59.

COLE, M. D. and COWLING, V. H. (2009). Specific regulation of mRNA cap methylation by the c-Myc and E2F1 transcription factors. *Oncogene,* 28**,** 1169-75.

COOPER, T. A., WAN, L. and DREYFUSS, G. (2009). RNA and disease. *Cell,* 136**,** 777-93.

COPPOLA, J. A., FIELD, A. S. and LUSE, D. S. (1983). Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc Natl Acad Sci U S A,* 80**,** 1251-5.

COWLING, V. H. (2010). Enhanced mRNA cap methylation increases cyclin D1 expression and promotes cell transformation. *Oncogene,* 29**,** 930-6.

COWLING, V. H. and COLE, M. D. (2007). The Myc transactivation domain promotes global phosphorylation of the RNA polymerase II carboxy-terminal domain independently of direct DNA binding. *Mol Cell Biol,* 27**,** 2059-73.

COWLING, V. H. and COLE, M. D. (2010). Myc Regulation of mRNA Cap Methylation. *Genes Cancer,* 1**,** 576-579.

CRICK, F. (1970). Central dogma of molecular biology. *Nature,* 227**,** 561-3.

CULJKOVIC, B., TOPISIROVIC, I., SKRABANEK, L., RUIZ-GUTIERREZ, M. and BORDEN, K. L. (2006). eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J Cell Biol,* 175**,** 415-26.

DARNELL, J. E., JR. (1997). STATs and gene regulation. *Science,* 277**,** 1630-5.

DARNELL, R. B. (2010). HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip Rev RNA,* 1**,** 266-86.

DAS, S., JENA, S. and LEVASSEUR, D. N. (2011). Alternative splicing produces Nanog protein variants with different capacities for self-renewal and pluripotency in embryonic stem cells. *J Biol Chem,* 286**,** 42690-703.

DAVIS, A. C., WIMS, M., SPOTTS, G. D., HANN, S. R. and BRADLEY, A. (1993). A null c-myc mutation causes lethality before 10.5 days of gestation in homozygotes and reduced fertility in heterozygous female mice. *Genes Dev,* 7**,** 671-82.

DELGADO, M. D. and LEON, J. (2006). Gene expression regulation and cancer. *Clin Transl Oncol,* 8**,** 780-7.

DESHAIES, R. J. and JOAZEIRO, C. A. (2009). RING domain E3 ubiquitin ligases. *Annu Rev Biochem,* 78**,** 399-434.

DI GIAMMARTINO, D. C., NISHIDA, K. and MANLEY, J. L. (2011). Mechanisms and consequences of alternative polyadenylation. *Mol Cell,* 43**,** 853-66.

DING, L., PASZKOWSKI-ROGACZ, M., NITZSCHE, A., SLABICKI, M. M., HENINGER, A. K., DE VRIES, I., KITTLER, R., JUNQUEIRA, M., SHEVCHENKO, A., SCHULZ, H., HUBNER, N., DOSS, M. X., SACHINIDIS, A., HESCHELER, J., IACONE, R., ANASTASSIADIS, K., STEWART, A. F., PISABARRO, M. T., CALDARELLI, A., POSER, I.,

THEIS, M. and BUCHHOLZ, F. (2009). A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell,* 4**,** 403-15.

DINGAR, D., KALKAT, M., CHAN, P. K., SRIKUMAR, T., BAILEY, S. D., TU, W. B., COYAUD, E., PONZIELLI, R., KOLYAR, M., JURISICA, I., HUANG, A., LUPIEN, M., PENN, L. Z. and RAUGHT, B. (2015). BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *J Proteomics,* 118**,** 95-111.

DRUMMOND, D. R., ARMSTRONG, J. and COLMAN, A. (1985). The effect of capping and polyadenylation on the stability, movement and translation of synthetic messenger RNAs in Xenopus oocytes. *Nucleic Acids Res,* 13**,** 7375-94.

DUNN, S. and COWLING, V. H. (2015). Myc and mRNA capping. *Biochimica et biophysica acta,* 1849**,** 501-505.

EDERY, I. and SONENBERG, N. (1985). Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc Natl Acad Sci U S A,* 82**,** 7590-4.

EFRONI, S., DUTTAGUPTA, R., CHENG, J., DEHGHANI, H., HOEPPNER, D. J., DASH, C., BAZETT-JONES, D. P., LE GRICE, S., MCKAY, R. D., BUETOW, K. H., GINGERAS, T. R., MISTELI, T. and MESHORER, E. (2008). Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell,* 2**,** 437-47.

EGLOFF, S. and MURPHY, S. (2008). Cracking the RNA polymerase II CTD code. *Trends Genet,* 24**,** 280-8.

EVANS, M. J. and KAUFMAN, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292**,** 154-6.

FABREGA, C., HAUSMANN, S., SHEN, V., SHUMAN, S. and LIMA, C. D. (2004). Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. *Mol Cell,* 13**,** 77-89.

FABREGA, C., SHEN, V., SHUMAN, S. and LIMA, C. D. (2003). Structure of an mRNA Capping Enzyme Bound to the Phosphorylated Carboxy-Terminal Domain of RNA Polymerase II. *Molecular Cell,* 11**,** 1549-1561.

FAZZIO, T. G., HUFF, J. T. and PANNING, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell,* 134**,** 162-74.

FEAVER, W. J., GILEADI, O., LI, Y. and KORNBERG, R. D. (1991). CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell,* 67**,** 1223-30.

FERNANDEZ-SANCHEZ, M. E., GONATOPOULOS-POURNATZIS, T., PRESTON, G., LAWLOR, M. A. and COWLING, V. H. (2009). S-adenosyl homocysteine hydrolase is required for Myc-induced mRNA cap methylation, protein synthesis, and cell proliferation. *Molecular and cellular biology,* 29**,** 6182-91.

FIDALGO, M., FAIOLA, F., PEREIRA, C. F., DING, J., SAUNDERS, A., GINGOLD, J., SCHANIEL, C., LEMISCHKA, I. R., SILVA, J. C. and WANG, J. (2012). Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proc Natl Acad Sci U S A,* 109**,** 16202-7.

FLAHERTY, S. M., FORTES, P., IZAURRALDE, E., MATTAJ, I. W. and GILMARTIN, G. M. (1997). Participation of the nuclear cap binding

complex in pre-mRNA 3' processing. *Proc Natl Acad Sci U S A,* 94**,** 11893-8.
FRANKE, W. W., GRUND, C., KUHN, C., JACKSON, B. W. and ILLMENSEE, K. (1982). Formation of cytoskeletal elements during mouse embryogenesis. III. Primary mesenchymal cells and the first appearance of vimentin filaments. *Differentiation,* 23**,** 43-59.
FRIED, H. and KUTAY, U. (2003). Nucleocytoplasmic transport: taking an inventory. *Cell Mol Life Sci,* 60**,** 1659-88.
FU, X. D. and ARES, M., JR. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. *Nat Rev Genet,* 15**,** 689-701.
FURUICHI, Y., LAFIANDRA, A. and SHATKIN, A. J. (1977). 5'-Terminal structure and mRNA stability. *Nature,* 266**,** 235-9.
FURUICHI, Y. and MIURA, K. (1975). A blocked structure at the 5' terminus of mRNA from cytoplasmic polyhedrosis virus. *Nature,* 253**,** 374-5.
FURUICHI, Y., MORGAN, M., MUTHUKRISHNAN, S. and SHATKIN, A. J. (1975a). Reovirus messenger RNA contains a methylated, blocked 5'-terminal structure: m-7G(5')ppp(5')G-MpCp. *Proc Natl Acad Sci U S A,* 72**,** 362-6.
FURUICHI, Y., MORGAN, M., SHATKIN, A. J., JELINEK, W., SALDITT-GEORGIEFF, M. and DARNELL, J. E. (1975b). Methylated, blocked 5 termini in HeLa cell mRNA. *Proc Natl Acad Sci U S A,* 72**,** 1904-8.
FURUICHI, Y. and SHATKIN, A. J. (2000). Viral and cellular mRNA capping: past and prospects. *Adv Virus Res,* 55**,** 135-84.
GABUT, M., SAMAVARCHI-TEHRANI, P., WANG, X., SLOBODENIUC, V., O'HANLON, D., SUNG, H. K., ALVAREZ, M., TALUKDER, S., PAN, Q., MAZZONI, E. O., NEDELEC, S., WICHTERLE, H., WOLTJEN, K., HUGHES, T. R., ZANDSTRA, P. W., NAGY, A., WRANA, J. L. and BLENCOWE, B. J. (2011). An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell,* 147**,** 132-46.
GAO, M., FRITZ, D. T., FORD, L. P. and WILUSZ, J. (2000). Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro. *Mol Cell,* 5**,** 479-88.
GEORGIEV, O., MOUS, J. and BIRNSTIEL, M. L. (1984). Processing and nucleo-cytoplasmic transport of histone gene transcripts. *Nucleic Acids Res,* 12**,** 8539-51.
GERRARD, L., RODGERS, L. and CUI, W. (2005). Differentiation of human embryonic stem cells to neural lineages in adherent culture by blocking bone morphogenetic protein signaling. *Stem Cells,* 23**,** 1234-41.
GEULA, S., MOSHITCH-MOSHKOVITZ, S., DOMINISSINI, D., MANSOUR, A. A., KOL, N., SALMON-DIVON, M., HERSHKOVITZ, V., PEER, E., MOR, N., MANOR, Y. S., BEN-HAIM, M. S., EYAL, E., YUNGER, S., PINTO, Y., JAITIN, D. A., VIUKOV, S., RAIS, Y., KRUPALNIK, V., CHOMSKY, E., ZERBIB, M., MAZA, I., RECHAVI, Y., MASSARWA, R., HANNA, S., AMIT, I., LEVANON, E. Y., AMARIGLIO, N., STERN-GINOSSAR, N., NOVERSHTERN, N., RECHAVI, G. and HANNA, J. H. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science,* 347**,** 1002-6.
GILLIAN-DANIEL, D. L., GRAY, N. K., ASTROM, J., BARKOFF, A. and WICKENS, M. (1998). Modifications of the 5' cap of mRNAs during

Xenopus oocyte maturation: independence from changes in poly(A) length and impact on translation. *Mol Cell Biol,* 18**,** 6152-63.
GLOVER-CUTTER, K., KIM, S., ESPINOSA, J. and BENTLEY, D. L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat Struct Mol Biol,* 15**,** 71-8.
GONATOPOULOS-POURNATZIS, T. and COWLING, V. H. (2014). RAM function is dependent on Kapbeta2-mediated nuclear entry. *Biochem J,* 457**,** 473-84.
GONATOPOULOS-POURNATZIS, T., DUNN, S., BOUNDS, R. and COWLING, V. H. (2011). RAM/Fam103a1 is required for mRNA cap methylation. *Molecular cell,* 44**,** 585-96.
GOSWAMI, T., LI, X., SMITH, A. M., LUDEROWSKI, E. M., VINCENT, J. J., RUSH, J. and BALLIF, B. A. (2012). Comparative phosphoproteomic analysis of neonatal and adult murine brain. *Proteomics,* 12**,** 2185-9.
GREEN, M. R., MANIATIS, T. and MELTON, D. A. (1983). Human beta-globin pre-mRNA synthesized in vitro is accurately spliced in Xenopus oocyte nuclei. *Cell,* 32**,** 681-94.
GROSS, J. D., MOERKE, N. J., VON DER HAAR, T., LUGOVSKOY, A. A., SACHS, A. B., MCCARTHY, J. E. and WAGNER, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell,* 115**,** 739-50.
GRUDZIEN, E., KALEK, M., JEMIELITY, J., DARZYNKIEWICZ, E. and RHOADS, R. E. (2006). Differential inhibition of mRNA degradation pathways by novel cap analogs. *J Biol Chem,* 281**,** 1857-67.
GU, M., RAJASHANKAR, K. R. and LIMA, C. D. (2010). Structure of the Saccharomyces cerevisiae Cet1-Ceg1 mRNA capping apparatus. *Structure,* 18**,** 216-27.
GUALLAR, D. and WANG, J. (2014). RNA-binding proteins in pluripotency, differentiation, and reprogramming. *Frontiers in biology,* 9**,** 389-409.
GUENTHER, M. G., LEVINE, S. S., BOYER, L. A., JAENISCH, R. and YOUNG, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell,* 130**,** 77-88.
GUIGUEN, A., SOUTOURINA, J., DEWEZ, M., TAFFOREAU, L., DIEU, M., RAES, M., VANDENHAUTE, J., WERNER, M. and HERMAND, D. (2007). Recruitment of P-TEFb (Cdk9-Pch1) to chromatin by the cap-methyl transferase Pcm1 in fission yeast. *EMBO J,* 26**,** 1552-9.
HAGHIGHAT, A. and SONENBERG, N. (1997). eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J Biol Chem,* 272**,** 21677-80.
HAHN, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol,* 11**,** 394-403.
HAKANSSON, K., DOHERTY, A. J., SHUMAN, S. and WIGLEY, D. B. (1997). X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell,* 89**,** 545-53.
HAKANSSON, K. and WIGLEY, D. B. (1998). Structure of a complex between a cap analogue and mRNA guanylyl transferase demonstrates the structural chemistry of RNA capping. *Proc Natl Acad Sci U S A,* 95**,** 1505-10.

HAMILTON, W. B. and BRICKMAN, J. M. (2014). Erk signaling suppresses embryonic stem cell self-renewal to specify endoderm. *Cell Rep,* 9**,** 2056-70.
HEIDEMANN, M., HINTERMAIR, C., VOSS, K. and EICK, D. (2013). Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta,* 1829**,** 55-62.
HNISZ, D., ABRAHAM, B. J., LEE, T. I., LAU, A., SAINT-ANDRE, V., SIGOVA, A. A., HOKE, H. A. and YOUNG, R. A. (2013). Super-enhancers in the control of cell identity and disease. *Cell,* 155**,** 934-47.
HO, C. K. and SHUMAN, S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell,* 3**,** 405-11.
HO, C. K., SRISKANDA, V., MCCRACKEN, S., BENTLEY, D., SCHWER, B. and SHUMAN, S. (1998). The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem,* 273**,** 9577-85.
HOCINE, S., SINGER, R. H. and GRUNWALD, D. (2010). RNA processing and export. *Cold Spring Harb Perspect Biol,* 2**,** a000752.
HOSKINS, A. A. and MOORE, M. J. (2012). The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem Sci,* 37**,** 179-88.
HSIN, J. P. and MANLEY, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev,* 26**,** 2119-37.
HU, G., KIM, J., XU, Q., LENG, Y., ORKIN, S. H. and ELLEDGE, S. J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev,* 23**,** 837-48.
HUNTER, T. (2007). The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol Cell,* 28**,** 730-8.
INGOLIA, N. T., LAREAU, L. F. and WEISSMAN, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell,* 147**,** 789-802.
INOUE, K., OHNO, M., SAKAMOTO, H. and SHIMURA, Y. (1989). Effect of the cap structure on pre-mRNA splicing in Xenopus oocyte nuclei. *Genes Dev,* 3**,** 1472-9.
IZAURRALDE, E., LEWIS, J., GAMBERI, C., JARMOLOWSKI, A., MCGUIGAN, C. and MATTAJ, I. W. (1995a). A cap-binding protein complex mediating U snRNA export. *Nature,* 376**,** 709-12.
IZAURRALDE, E., LEWIS, J., MCGUIGAN, C., JANKOWSKA, M., DARZYNKIEWICZ, E. and MATTAJ, I. W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell,* 78**,** 657-68.
IZAURRALDE, E., MCGUIGAN, C. and MATTAJ, I. W. (1995b). Nuclear localization of a cap-binding protein complex. *Cold Spring Harb Symp Quant Biol,* 60**,** 669-75.
JACKSON, R. J., HELLEN, C. U. and PESTOVA, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol,* 11**,** 113-27.
JAENISCH, R. and YOUNG, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell,* 132**,** 567-82.
JIA, J., ZHENG, X., HU, G., CUI, K., ZHANG, J., ZHANG, A., JIANG, H., LU, B., YATES, J., 3RD, LIU, C., ZHAO, K. and ZHENG, Y. (2012). Regulation

of pluripotency and self- renewal of ESCs through epigenetic-threshold modulation and mRNA pruning. *Cell,* 151**,** 576-89.

JIAO, X., CHANG, J. H., KILIC, T., TONG, L. and KILEDJIAN, M. (2013). A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol Cell,* 50**,** 104-15.

JIAO, X., XIANG, S., OH, C., MARTIN, C. E., TONG, L. and KILEDJIAN, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature,* 467**,** 608-11.

JOVE, R. and MANLEY, J. L. (1984). In vitro transcription from the adenovirus 2 major late promoter utilizing templates truncated at promoter-proximal sites. *J Biol Chem,* 259**,** 8513-21.

KAGEY, M. H., NEWMAN, J. J., BILODEAU, S., ZHAN, Y., ORLANDO, D. A., VAN BERKUM, N. L., EBMEIER, C. C., GOOSSENS, J., RAHL, P. B., LEVINE, S. S., TAATJES, D. J., DEKKER, J. and YOUNG, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature,* 467**,** 430-5.

KAHVEJIAN, A., SVITKIN, Y. V., SUKARIEH, R., M'BOUTCHOU, M. N. and SONENBERG, N. (2005). Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev,* 19**,** 104-13.

KATAHIRA, J. (2012). mRNA export and the TREX complex. *Biochim Biophys Acta,* 1819**,** 507-13.

KATAOKA, N., OHNO, M., MODA, I. and SHIMURA, Y. (1995). Identification of the factors that interact with NCBP, an 80 kDa nuclear cap binding protein. *Nucleic Acids Res,* 23**,** 3638-41.

KHANNA, R. and KILEDJIAN, M. (2004). Poly(A)-binding-protein-mediated regulation of hDcp2 decapping in vitro. *EMBO J,* 23**,** 1968-76.

KIM, J., CHU, J., SHEN, X., WANG, J. and ORKIN, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell,* 132**,** 1049-61.

KIM, T. H., BARRERA, L. O., ZHENG, M., QU, C., SINGER, M. A., RICHMOND, T. A., WU, Y., GREEN, R. D. and REN, B. (2005). A high-resolution map of active promoters in the human genome. *Nature,* 436**,** 876-80.

KLEINSMITH, L. J. and PIERCE, G. B., JR. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. *Cancer Res,* 24**,** 1544-51.

KOHLER, A. and HURT, E. (2007). Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol,* 8**,** 761-73.

KOMAR, A. A. and HATZOGLOU, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell Cycle,* 10**,** 229-40.

KOMARNITSKY, P., CHO, E. J. and BURATOWSKI, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev,* 14**,** 2452-60.

KONARSKA, M. M., PADGETT, R. A. and SHARP, P. A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell,* 38**,** 731-6.

KOROMILAS, A. E., LAZARIS-KARATZAS, A. and SONENBERG, N. (1992). mRNAs containing extensive secondary structure in their 5' non-coding

region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J,* 11**,** 4153-8.

KRAVCHENKO, J. E., ROGOZIN, I. B., KOONIN, E. V. and CHUMAKOV, P. M. (2005). Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. *Nature,* 436**,** 735-9.

KU, M., KOCHE, R. P., RHEINBAY, E., MENDENHALL, E. M., ENDOH, M., MIKKELSEN, T. S., PRESSER, A., NUSBAUM, C., XIE, X., CHI, A. S., ADLI, M., KASIF, S., PTASZEK, L. M., COWAN, C. A., LANDER, E. S., KOSEKI, H. and BERNSTEIN, B. E. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet,* 4**,** e1000242.

KUNATH, T., SABA-EL-LEIL, M. K., ALMOUSAILLEAKH, M., WRAY, J., MELOCHE, S. and SMITH, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development,* 134**,** 2895-902.

KWON, S. C., YI, H., EICHELBAUM, K., FOHR, S., FISCHER, B., YOU, K. T., CASTELLO, A., KRIJGSVELD, J., HENTZE, M. W. and KIM, V. N. (2013). The RNA-binding protein repertoire of embryonic stem cells. *Nature structural & molecular biology,* 20**,** 1122-30.

LACKFORD, B., YAO, C., CHARLES, G. M., WENG, L., ZHENG, X., CHOI, E. A., XIE, X., WAN, J., XING, Y., FREUDENBERG, J. M., YANG, P., JOTHI, R., HU, G. and SHI, Y. (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J,* 33**,** 878-89.

LAHUDKAR, S., SHUKLA, A., BAJWA, P., DURAIRAJ, G., STANOJEVIC, N. and BHAUMIK, S. R. (2011). The mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p in vivo. *Nucleic Acids Res,* 39**,** 2188-209.

LANGE, A., MILLS, R. E., LANGE, C. J., STEWART, M., DEVINE, S. E. and CORBETT, A. H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem,* 282**,** 5101-5.

LAZARIS-KARATZAS, A., MONTINE, K. S. and SONENBERG, N. (1990). Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature,* 345**,** 544-7.

LEE, T. I., JENNER, R. G., BOYER, L. A., GUENTHER, M. G., LEVINE, S. S., KUMAR, R. M., CHEVALIER, B., JOHNSTONE, S. E., COLE, M. F., ISONO, K., KOSEKI, H., FUCHIKAMI, T., ABE, K., MURRAY, H. L., ZUCKER, J. P., YUAN, B., BELL, G. W., HERBOLSHEIMER, E., HANNETT, N. M., SUN, K., ODOM, D. T., OTTE, A. P., VOLKERT, T. L., BARTEL, D. P., MELTON, D. A., GIFFORD, D. K., JAENISCH, R. and YOUNG, R. A. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell,* 125**,** 301-13.

LENASI, T., PETERLIN, B. M. and BARBORIC, M. (2011). Cap-binding protein complex links pre-mRNA capping to transcription elongation and alternative splicing through positive transcription elongation factor b (P-TEFb). *J Biol Chem,* 286**,** 22758-68.

LENDAHL, U., ZIMMERMAN, L. B. and MCKAY, R. D. (1990). CNS stem cells express a new class of intermediate filament protein. *Cell,* 60**,** 585-95.

LEVINE, M. and TJIAN, R. (2003). Transcription regulation and animal diversity. *Nature,* 424**,** 147-51.

LI, Y., SONG, M. and KILEDJIAN, M. (2011). Differential utilization of decapping enzymes in mammalian mRNA decay pathways. *RNA,* 17**,** 419-28.
LIMA, C. D., WANG, L. K. and SHUMAN, S. (1999). Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell,* 99**,** 533-43.
LIN, C. Y., LOVEN, J., RAHL, P. B., PARANAL, R. M., BURGE, C. B., BRADNER, J. E., LEE, T. I. and YOUNG, R. A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. *Cell,* 151**,** 56-67.
LISTERMAN, I., SAPRA, A. K. and NEUGEBAUER, K. M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol,* 13**,** 815-22.
LIU, H., RODGERS, N. D., JIAO, X. and KILEDJIAN, M. (2002). The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J,* 21**,** 4699-708.
LIU, P., KENNEY, J. M., STILLER, J. W. and GREENLEAF, A. L. (2010). Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain. *Mol Biol Evol,* 27**,** 2628-41.
LIU, Y., LEE, M. R., TIMANI, K., HE, J. J. and BROXMEYER, H. E. (2012). Tip110 maintains expression of pluripotent factors in and pluripotency of human embryonic stem cells. *Stem Cells Dev,* 21**,** 829-33.
LIU, Y., TIMANI, K., OU, X., BROXMEYER, H. E. and HE, J. J. (2013). C-MYC controlled TIP110 protein expression regulates OCT4 mRNA splicing in human embryonic stem cells. *Stem Cells Dev,* 22**,** 689-94.
LOH, Y. H., WU, Q., CHEW, J. L., VEGA, V. B., ZHANG, W., CHEN, X., BOURQUE, G., GEORGE, J., LEONG, B., LIU, J., WONG, K. Y., SUNG, K. W., LEE, C. W., ZHAO, X. D., CHIU, K. P., LIPOVICH, L., KUZNETSOV, V. A., ROBSON, P., STANTON, L. W., WEI, C. L., RUAN, Y., LIM, B. and NG, H. H. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet,* 38**,** 431-40.
LU, H., FLORES, O., WEINMANN, R. and REINBERG, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc Natl Acad Sci U S A,* 88**,** 10004-8.
LU, X., GOKE, J., SACHS, F., JACQUES, P. E., LIANG, H., FENG, B., BOURQUE, G., BUBULYA, P. A. and NG, H. H. (2013). SON connects the splicing-regulatory network with pluripotency in human embryonic stem cells. *Nat Cell Biol,* 15**,** 1141-52.
LYKKE-ANDERSEN, J. (2002). Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol,* 22**,** 8114-21.
MANDAL, S. S., CHU, C., WADA, T., HANDA, H., SHATKIN, A. J. and REINBERG, D. (2004). Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proceedings of the National Academy of Sciences,* 101**,** 7572-7577.
MAO, X., SCHWER, B. and SHUMAN, S. (1995). Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene. *Mol Cell Biol,* 15**,** 4167-74.

MARCHETTO, M. C., YEO, G. W., KAINOHANA, O., MARSALA, M., GAGE, F. H. and MUOTRI, A. R. (2009). Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS One,* 4**,** e7076.

MARCOTRIGIANO, J., GINGRAS, A. C., SONENBERG, N. and BURLEY, S. K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell,* 89**,** 951-61.

MARTI, M., MULERO, L., PARDO, C., MORERA, C., CARRIO, M., LARICCHIA-ROBBIO, L., ESTEBAN, C. R. and IZPISUA BELMONTE, J. C. (2013). Characterization of pluripotent stem cells. *Nat Protoc,* 8**,** 223-53.

MARTIN, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A,* 78**,** 7634-8.

MARTIN, G. R. and EVANS, M. J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proc Natl Acad Sci U S A,* 72**,** 1441-5.

MARTINEZ-RUCOBO, F. W., KOHLER, R., VAN DE WATERBEEMD, M., HECK, A. J., HEMANN, M., HERZOG, F., STARK, H. and CRAMER, P. (2015). Molecular Basis of Transcription-Coupled Pre-mRNA Capping. *Mol Cell,* 58**,** 1079-89.

MASEK, T., VALASEK, L. and POSPISEK, M. (2011). Polysome analysis and RNA purification from sucrose gradients. *Methods Mol Biol,* 703**,** 293-309.

MASUI, S., NAKATAKE, Y., TOYOOKA, Y., SHIMOSATO, D., YAGI, R., TAKAHASHI, K., OKOCHI, H., OKUDA, A., MATOBA, R., SHAROV, A. A., KO, M. S. and NIWA, H. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol,* 9**,** 625-35.

MATHUR, D., DANFORD, T. W., BOYER, L. A., YOUNG, R. A., GIFFORD, D. K. and JAENISCH, R. (2008). Analysis of the mouse embryonic stem cell regulatory networks obtained by ChIP-chip and ChIP-PET. *Genome Biol,* 9**,** R126.

MAZZA, C., OHNO, M., SEGREF, A., MATTAJ, I. W. and CUSACK, S. (2001). Crystal structure of the human nuclear cap binding complex. *Mol Cell,* 8**,** 383-96.

MAZZA, C., SEGREF, A., MATTAJ, I. W. and CUSACK, S. (2002). Co-crystallization of the human nuclear cap-binding complex with a m7GpppG cap analogue using protein engineering. *Acta Crystallogr D Biol Crystallogr,* 58**,** 2194-7.

MCCLOSKEY, A., TANIGUCHI, I., SHINMYOZU, K. and OHNO, M. (2012). hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science,* 335**,** 1643-6.

MCCRACKEN, S., FONG, N., ROSONINA, E., YANKULOV, K., BROTHERS, G., SIDEROVSKI, D., HESSEL, A., FOSTER, S., SHUMAN, S. and BENTLEY, D. L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev,* 11**,** 3306-18.

MESHORER, E. and MISTELI, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol,* 7**,** 540-6.

MESHORER, E., YELLAJOSHULA, D., GEORGE, E., SCAMBLER, P. J., BROWN, D. T. and MISTELI, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell,* 10**,** 105-16.

MEYER, N. and PENN, L. Z. (2008). Reflecting on 25 years with MYC. *Nat Rev Cancer,* 8**,** 976-90.

MIKKELSEN, T. S., KU, M., JAFFE, D. B., ISSAC, B., LIEBERMAN, E., GIANNOUKOS, G., ALVAREZ, P., BROCKMAN, W., KIM, T. K., KOCHE, R. P., LEE, W., MENDENHALL, E., O'DONOVAN, A., PRESSER, A., RUSS, C., XIE, X., MEISSNER, A., WERNIG, M., JAENISCH, R., NUSBAUM, C., LANDER, E. S. and BERNSTEIN, B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature,* 448**,** 553-60.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. and YAMANAKA, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* 113**,** 631-42.

MOORE, M. J. and PROUDFOOT, N. J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell,* 136**,** 688-700.

MORRISON, S. J. and KIMBLE, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature,* 441**,** 1068-74.

MOTEKI, S. and PRICE, D. (2002). Functional coupling of capping and transcription of mRNA. *Mol Cell,* 10**,** 599-609.

MUELLER, A. A., CHEUNG, T. H. and RANDO, T. A. (2013). All's well that ends well: alternative polyadenylation and its implications for stem cell biology. *Curr Opin Cell Biol,* 25**,** 222-32.

MURTHY, K. G., PARK, P. and MANLEY, J. L. (1991). A nuclear micrococcal-sensitive, ATP-dependent exoribonuclease degrades uncapped but not capped RNA substrates. *Nucleic Acids Res,* 19**,** 2685-92.

MUTHUKRISHNAN, S., BOTH, G. W., FURUICHI, Y. and SHATKIN, A. J. (1975). 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature,* 255**,** 33-7.

NARITA, T., YUNG, T. M., YAMAMOTO, J., TSUBOI, Y., TANABE, H., TANAKA, K., YAMAGUCHI, Y. and HANDA, H. (2007). NELF interacts with CBC and participates in 3' end processing of replication-dependent histone mRNAs. *Mol Cell,* 26**,** 349-65.

NAVARRO, P., FESTUCCIA, N., COLBY, D., GAGLIARDI, A., MULLIN, N. P., ZHANG, W., KARWACKI-NEISIUS, V., OSORNO, R., KELLY, D., ROBERTSON, M. and CHAMBERS, I. (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *EMBO J,* 31**,** 4547-62.

NICHOLS, J., ZEVNIK, B., ANASTASSIADIS, K., NIWA, H., KLEWE-NEBENIUS, D., CHAMBERS, I., SCHOLER, H. and SMITH, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell,* 95**,** 379-91.

NIE, Z., HU, G., WEI, G., CUI, K., YAMANE, A., RESCH, W., WANG, R., GREEN, D. R., TESSAROLLO, L., CASELLAS, R., ZHAO, K. and LEVENS, D. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell,* 151**,** 68-79.

NIKIFOROV, M. A., CHANDRIANI, S., PARK, J., KOTENKO, I., MATHEOS, D., JOHNSSON, A., MCMAHON, S. B. and COLE, M. D. (2002). TRRAP-dependent and TRRAP-independent transcriptional activation by Myc family oncoproteins. *Mol Cell Biol,* 22**,** 5054-63.

NIKOLOV, D. B. and BURLEY, S. K. (1997). RNA polymerase II transcription initiation: a structural view. *Proc Natl Acad Sci U S A,* 94**,** 15-22.

NISHIMOTO, M., FUKUSHIMA, A., OKUDA, A. and MURAMATSU, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol Cell Biol,* 19**,** 5453-65.

NIWA, H., BURDON, T., CHAMBERS, I. and SMITH, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev,* 12**,** 2048-60.

NIWA, H., MIYAZAKI, J. and SMITH, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet,* 24**,** 372-6.

OHNO, M., SAKAMOTO, H. and SHIMURA, Y. (1987). Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure. *Proc Natl Acad Sci U S A,* 84**,** 5187-91.

OHTSUKA, S. and DALTON, S. (2008). Molecular and biological properties of pluripotent embryonic stem cells. *Gene Ther,* 15**,** 74-81.

ORPHANIDES, G. and REINBERG, D. (2002). A unified theory of gene expression. *Cell,* 108**,** 439-51.

PABIS, M., NEUFELD, N., STEINER, M. C., BOJIC, T., SHAV-TAL, Y. and NEUGEBAUER, K. M. (2013). The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA,* 19**,** 1054-63.

PATZELT, E., THALMANN, E., HARTMUTH, K., BLAAS, D. and KUECHLER, E. (1987). Assembly of pre-mRNA splicing complex is cap dependent. *Nucleic Acids Res,* 15**,** 1387-99.

PEI, Y. and SHUMAN, S. (2002). Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *J Biol Chem,* 277**,** 19639-48.

PETERLIN, B. M. and PRICE, D. H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Mol Cell,* 23**,** 297-305.

PEUKERT, K., STALLER, P., SCHNEIDER, A., CARMICHAEL, G., HANEL, F. and EILERS, M. (1997). An alternative pathway for gene regulation by Myc. *EMBO J,* 16**,** 5672-86.

PEVNY, L. H., SOCKANATHAN, S., PLACZEK, M. and LOVELL-BADGE, R. (1998). A role for SOX1 in neural determination. *Development,* 125**,** 1967-78.

PHILLIPS, K. and LUISI, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. *J Mol Biol,* 302**,** 1023-39.

PIJNAPPEL, W. W., ESCH, D., BALTISSEN, M. P., WU, G., MISCHERIKOW, N., BERGSMA, A. J., VAN DER WAL, E., HAN, D. W., BRUCH, H., MORITZ, S., LIJNZAAD, P., ALTELAAR, A. F., SAMEITH, K., ZAEHRES, H., HECK, A. J., HOLSTEGE, F. C., SCHOLER, H. R. and TIMMERS, H. T. (2013). A central role for TFIID in the pluripotent transcription circuitry. *Nature,* 495**,** 516-9.

PILLUTLA, R. C., SHIMAMOTO, A., FURUICHI, Y. and SHATKIN, A. J. (1998a). Human mRNA capping enzyme (RNGTT) and cap methyltransferase (RNMT) map to 6q16 and 18p11.22-p11.23, respectively. *Genomics,* 54**,** 351-3.

PILLUTLA, R. C., YUE, Z., MALDONADO, E. and SHATKIN, A. J. (1998b). Recombinant human mRNA cap methyltransferase binds capping enzyme/RNA polymerase IIo complexes. *J Biol Chem,* 273**,** 21443-6.

RAHL, P. B., LIN, C. Y., SEILA, A. C., FLYNN, R. A., MCCUINE, S., BURGE, C. B., SHARP, P. A. and YOUNG, R. A. (2010). c-Myc regulates transcriptional pause release. *Cell,* 141**,** 432-45.

RAMOS-MEJIA, V., MUNOZ-LOPEZ, M., GARCIA-PEREZ, J. L. and MENENDEZ, P. (2010). iPSC lines that do not silence the expression of the ectopic reprogramming factors may display enhanced propensity to genomic instability. *Cell Res,* 20**,** 1092-5.

RAO, S., ZHEN, S., ROUMIANTSEV, S., MCDONALD, L. T., YUAN, G. C. and ORKIN, S. H. (2010). Differential roles of Sall4 isoforms in embryonic stem cell pluripotency. *Mol Cell Biol,* 30**,** 5364-80.

RASMUSSEN, E. B. and LIS, J. T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. *Proc Natl Acad Sci U S A,* 90**,** 7923-7.

REDDY, R., SINGH, R. and SHIMBA, S. (1992). Methylated cap structures in eukaryotic RNAs: structure, synthesis and functions. *Pharmacol Ther,* 54**,** 249-67.

ROEDER, R. G. and RUTTER, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature,* 224**,** 234-7.

RUGGERO, D., MONTANARO, L., MA, L., XU, W., LONDEI, P., CORDON-CARDO, C. and PANDOLFI, P. P. (2004). The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nat Med,* 10**,** 484-6.

SAHA, N., SCHWER, B. and SHUMAN, S. (1999). Characterization of human, Schizosaccharomyces pombe, and Candida albicans mRNA cap methyltransferases and complete replacement of the yeast capping apparatus by mammalian enzymes. *J Biol Chem,* 274**,** 16553-62.

SAMPATH, P., PRITCHARD, D. K., PABON, L., REINECKE, H., SCHWARTZ, S. M., MORRIS, D. R. and MURRY, C. E. (2008). A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell,* 2**,** 448-60.

SATO, H. and MAQUAT, L. E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. *Genes Dev,* 23**,** 2537-50.

SATO, N., MEIJER, L., SKALTSOUNIS, L., GREENGARD, P. and BRIVANLOU, A. H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med,* 10**,** 55-63.

SCHNEIDER-POETSCH, T., JU, J., EYLER, D. E., DANG, Y., BHAT, S., MERRICK, W. C., GREEN, R., SHEN, B. and LIU, J. O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol,* 6**,** 209-217.

SCHROEDER, S. C., SCHWER, B., SHUMAN, S. and BENTLEY, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev,* 14**,** 2435-40.

SCHWARTZ, D. C. and PARKER, R. (2000). mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Mol Cell Biol,* 20**,** 7933-42.

SCHWER, B., SAHA, N., MAO, X., CHEN, H. W. and SHUMAN, S. (2000). Structure-function analysis of yeast mRNA cap methyltransferase and high-copy suppression of conditional mutants by AdoMet synthase and the ubiquitin conjugating enzyme Cdc34p. *Genetics,* 155**,** 1561-76.

SEO, J. and LEE, K. J. (2004). Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *J Biochem Mol Biol,* 37**,** 35-44.

SHAFER, B., CHU, C. and SHATKIN, A. J. (2005). Human mRNA cap methyltransferase: alternative nuclear localization signal motifs ensure nuclear localization required for viability. *Mol Cell Biol,* 25**,** 2644-9.

SHARP, P. A. (2009). The centrality of RNA. *Cell,* 136**,** 577-80.

SHATKIN, A. J. (1976). Capping of eucaryotic mRNAs. *Cell,* 9**,** 645-53.

SHATKIN, A. J. and MANLEY, J. L. (2000). The ends of the affair: capping and polyadenylation. *Nat Struct Biol,* 7**,** 838-42.

SHIBAGAKI, Y., ITOH, N., YAMADA, H., NAGATA, S. and MIZUMOTO, K. (1992). mRNA capping enzyme. Isolation and characterization of the gene encoding mRNA guanylytransferase subunit from Saccharomyces cerevisiae. *J Biol Chem,* 267**,** 9521-8.

SHIMOTOHNO, K., KODAMA, Y., HASHIMOTO, J. and MIURA, K. I. (1977). Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proc Natl Acad Sci U S A,* 74**,** 2734-8.

SHIRAKI, N., SHIRAKI, Y., TSUYAMA, T., OBATA, F., MIURA, M., NAGAE, G., ABURATANI, H., KUME, K., ENDO, F. and KUME, S. (2014). Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. *Cell Metab,* 19**,** 780-94.

SHUMAN, S. (2002). What messenger RNA capping tells us about eukaryotic evolution. *Nat Rev Mol Cell Biol,* 3**,** 619-25.

SHYH-CHANG, N., LOCASALE, J. W., LYSSIOTIS, C. A., ZHENG, Y., TEO, R. Y., RATANASIRINTRAWOOT, S., ZHANG, J., ONDER, T., UNTERNAEHRER, J. J., ZHU, H., ASARA, J. M., DALEY, G. Q. and CANTLEY, L. C. (2013). Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science,* 339**,** 222-6.

SIKORSKI, T. W. and BURATOWSKI, S. (2009). The basal initiation machinery: beyond the general transcription factors. *Curr Opin Cell Biol,* 21**,** 344-51.

SMITH, A. G. (2001). Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol,* 17**,** 435-62.

SMITH, A. G., HEATH, J. K., DONALDSON, D. D., WONG, G. G., MOREAU, J., STAHL, M. and ROGERS, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature,* 336**,** 688-90.

SOGAARD, T. M. and SVEJSTRUP, J. Q. (2007). Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J Biol Chem,* 282**,** 14113-20.

SOLTER, D., SKREB, N. and DAMJANOV, I. (1970). Extrauterine growth of mouse egg-cylinders results in malignant teratoma. *Nature,* 227**,** 503-4.

SONENBERG, N. and HINNEBUSCH, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell,* 136**,** 731-45.

SONENBERG, N., MORGAN, M. A., MERRICK, W. C. and SHATKIN, A. J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci U S A,* 75**,** 4843-7.

SONENBERG, N., RUPPRECHT, K. M., HECHT, S. M. and SHATKIN, A. J. (1979). Eukaryotic mRNA cap binding protein: purification by affinity chromatography on sepharose-coupled m7GDP. *Proc Natl Acad Sci U S A,* 76**,** 4345-9.

SONENBERG, N., TRACHSEL, H., HECHT, S. and SHATKIN, A. J. (1980). Differential stimulation of capped mRNA translation in vitro by cap binding protein. *Nature,* 285**,** 331-3.

SONG, M. G., BAIL, S. and KILEDJIAN, M. (2013). Multiple Nudix family proteins possess mRNA decapping activity. *RNA,* 19**,** 390-9.

SONG, M. G., LI, Y. and KILEDJIAN, M. (2010). Multiple mRNA decapping enzymes in mammalian cells. *Mol Cell,* 40**,** 423-32.

SPITZ, F. and FURLONG, E. E. (2012). Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet,* 13**,** 613-26.

SPRIGGS, K. A., BUSHELL, M. and WILLIS, A. E. (2010). Translational regulation of gene expression during conditions of cell stress. *Mol Cell,* 40**,** 228-37.

STAVRIDIS, M. P., LUNN, J. S., COLLINS, B. J. and STOREY, K. G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development,* 134**,** 2889-94.

STEVENS, L. C. (1970). The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol,* 21**,** 364-82.

STEWART, C. L., KASPAR, P., BRUNET, L. J., BHATT, H., GADI, I., KONTGEN, F. and ABBONDANZO, S. J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature,* 359**,** 76-9.

STRASSER, K., MASUDA, S., MASON, P., PFANNSTIEL, J., OPPIZZI, M., RODRIGUEZ-NAVARRO, S., RONDON, A. G., AGUILERA, A., STRUHL, K., REED, R. and HURT, E. (2002). TREX is a conserved complex coupling transcription with messenger RNA export. *Nature,* 417**,** 304-8.

TAHMASEBI, S., ALAIN, T., RAJASEKHAR, V. K., ZHANG, J. P., PRAGER-KHOUTORSKY, M., KHOUTORSKY, A., DOGAN, Y., GKOGKAS, C. G., PETROULAKIS, E., SYLVESTRE, A., GHORBANI, M., ASSADIAN, S., YAMANAKA, Y., VINAGOLU-BAUR, J. R., TEODORO, J. G., KIM, K., YANG, X. J. and SONENBERG, N. (2014). Multifaceted regulation of somatic cell reprogramming by mRNA translational control. *Cell Stem Cell,* 14**,** 606-16.

TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. and YAMANAKA, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell,* 131**,** 861-72.

TAKAHASHI, K. and YAMANAKA, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* 126**,** 663-76.

THOMSON, J. A., ITSKOVITZ-ELDOR, J., SHAPIRO, S. S., WAKNITZ, M. A., SWIERGIEL, J. J., MARSHALL, V. S. and JONES, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science,* 282**,** 1145-7.

TONG, M., LV, Z., LIU, L., ZHU, H., ZHENG, Q. Y., ZHAO, X. Y., LI, W., WU, Y. B., ZHANG, H. J., WU, H. J., LI, Z. K., ZENG, F., WANG, L., WANG, X. J., SHA, J. H. and ZHOU, Q. (2011). Mice generated from tetraploid complementation competent iPS cells show similar developmental features as those from ES cells but are prone to tumorigenesis. *Cell Res,* 21**,** 1634-7.

TOPISIROVIC, I., SVITKIN, Y. V., SONENBERG, N. and SHATKIN, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley interdisciplinary reviews. RNA,* 2**,** 277-98.

TORRES-PADILLA, M. E. and CHAMBERS, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development,* 141**,** 2173-81.

TSAI, S. C., CHANG, D. F., HONG, C. M., XIA, P., SENADHEERA, D., TRUMP, L., MISHRA, S. and LUTZKO, C. (2014). Induced overexpression of OCT4A in human embryonic stem cells increases cloning efficiency. *Am J Physiol Cell Physiol,* 306**,** C1108-18.

TSUKAMOTO, T., SHIBAGAKI, Y., IMAJOH-OHMI, S., MURAKOSHI, T., SUZUKI, M., NAKAMURA, A., GOTOH, H. and MIZUMOTO, K. (1997). Isolation and characterization of the yeast mRNA capping enzyme beta subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability. *Biochem Biophys Res Commun,* 239**,** 116-22.

ULE, J., JENSEN, K., MELE, A. and DARNELL, R. B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods,* 37**,** 376-86.

UY, R. and WOLD, F. (1977). Posttranslational covalent modification of proteins. *Science,* 198**,** 890-6.

VON DER HAAR, T., GROSS, J. D., WAGNER, G. and MCCARTHY, J. E. (2004). The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat Struct Mol Biol,* 11**,** 503-11.

WADA, T., TAKAGI, T., YAMAGUCHI, Y., FERDOUS, A., IMAI, T., HIROSE, S., SUGIMOTO, S., YANO, K., HARTZOG, G. A., WINSTON, F., BURATOWSKI, S. and HANDA, H. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev,* 12**,** 343-56.

WALTHER, C. and GRUSS, P. (1991). Pax-6, a murine paired box gene, is expressed in the developing CNS. *Development,* 113**,** 1435-49.

WANG, J., RAO, S., CHU, J., SHEN, X., LEVASSEUR, D. N., THEUNISSEN, T. W. and ORKIN, S. H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature,* 444**,** 364-8.

WANG, L., DU, Y., WARD, J. M., SHIMBO, T., LACKFORD, B., ZHENG, X., MIAO, Y. L., ZHOU, B., HAN, L., FARGO, D. C., JOTHI, R., WILLIAMS, C. J., WADE, P. A. and HU, G. (2014a). INO80 facilitates pluripotency

gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. *Cell Stem Cell,* 14**,** 575-91.

WANG, L., MIAO, Y. L., ZHENG, X., LACKFORD, B., ZHOU, B., HAN, L., YAO, C., WARD, J. M., BURKHOLDER, A., LIPCHINA, I., FARGO, D. C., HOCHEDLINGER, K., SHI, Y., WILLIAMS, C. J. and HU, G. (2013). The THO complex regulates pluripotency gene mRNA export and controls embryonic stem cell self-renewal and somatic cell reprogramming. *Cell stem cell,* 13**,** 676-90.

WANG, Y., LI, Y., TOTH, J. I., PETROSKI, M. D., ZHANG, Z. and ZHAO, J. C. (2014b). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol,* 16**,** 191-8.

WANG, Y. C., PETERSON, S. E. and LORING, J. F. (2014c). Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res,* 24**,** 143-60.

WANG, Z., JIAO, X., CARR-SCHMID, A. and KILEDJIAN, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proceedings of the National Academy of Sciences,* 99**,** 12663-12668.

WANG, Z. and KILEDJIAN, M. (2001). Functional link between the mammalian exosome and mRNA decapping. *Cell,* 107**,** 751-62.

WELLS, S. E., HILLNER, P. E., VALE, R. D. and SACHS, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell,* 2**,** 135-40.

WEN, Y. and SHATKIN, A. J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev,* 13**,** 1774-9.

WEN, Y. and SHATKIN, A. J. (2000). Cap methyltransferase selective binding and methylation of GpppG-RNA are stimulated by importin-alpha. *Genes Dev,* 14**,** 2944-9.

WERNER, M., PURTA, E., KAMINSKA, K. H., CYMERMAN, I. A., CAMPBELL, D. A., MITTRA, B., ZAMUDIO, J. R., STURM, N. R., JAWORSKI, J. and BUJNICKI, J. M. (2011). 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res,* 39**,** 4756-68.

WHYTE, W. A., ORLANDO, D. A., HNISZ, D., ABRAHAM, B. J., LIN, C. Y., KAGEY, M. H., RAHL, P. B., LEE, T. I. and YOUNG, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell,* 153**,** 307-19.

WILLIAMS, D. C., JR., CAI, M. and CLORE, G. M. (2004). Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex. *J Biol Chem,* 279**,** 1449-57.

WILLIAMSON, R., SCALES, T., CLARK, B. R., GIBB, G., REYNOLDS, C. H., KELLIE, S., BIRD, I. N., VARNDELL, I. M., SHEPPARD, P. W., EVERALL, I. and ANDERTON, B. H. (2002). Rapid tyrosine phosphorylation of neuronal proteins including tau and focal adhesion kinase in response to amyloid-beta peptide exposure: involvement of Src family protein kinases. *J Neurosci,* 22**,** 10-20.

WILSON-GRADY, J. T., HAAS, W. and GYGI, S. P. (2013). Quantitative comparison of the fasted and re-fed mouse liver phosphoproteomes using lower pH reductive dimethylation. *Methods,* 61**,** 277-86.

WOLTJEN, K., MICHAEL, I. P., MOHSENI, P., DESAI, R., MILEIKOVSKY, M., HAMALAINEN, R., COWLING, R., WANG, W., LIU, P., GERTSENSTEIN, M., KAJI, K., SUNG, H. K. and NAGY, A. (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature,* 458**,** 766-70.

WONG, C. M., QIU, H., HU, C., DONG, J. and HINNEBUSCH, A. G. (2007). Yeast cap binding complex impedes recruitment of cleavage factor IA to weak termination sites. *Mol Cell Biol,* 27**,** 6520-31.

WORCH, R., NIEDZWIECKA, A., STEPINSKI, J., MAZZA, C., JANKOWSKA-ANYSZKA, M., DARZYNKIEWICZ, E., CUSACK, S. and STOLARSKI, R. (2005). Specificity of recognition of mRNA 5' cap by human nuclear cap-binding complex. *RNA,* 11**,** 1355-63.

WRAY, J., KALKAN, T., GOMEZ-LOPEZ, S., ECKARDT, D., COOK, A., KEMLER, R. and SMITH, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol,* 13**,** 838-45.

WU, J. Q., HABEGGER, L., NOISA, P., SZEKELY, A., QIU, C., HUTCHISON, S., RAHA, D., EGHOLM, M., LIN, H., WEISSMAN, S., CUI, W., GERSTEIN, M. and SNYDER, M. (2010). Dynamic transcriptomes during neural differentiation of human embryonic stem cells revealed by short, long, and paired-end sequencing. *Proc Natl Acad Sci U S A,* 107**,** 5254-9.

YAMADA-OKABE, T., DOI, R., SHIMMI, O., ARISAWA, M. and YAMADA-OKABE, H. (1998). Isolation and characterization of a human cDNA for mRNA 5'-capping enzyme. *Nucleic Acids Res,* 26**,** 1700-6.

YAMAGUCHI, Y., TAKAGI, T., WADA, T., YANO, K., FURUYA, A., SUGIMOTO, S., HASEGAWA, J. and HANDA, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell,* 97**,** 41-51.

YASUHARA, N., YAMAGISHI, R., ARAI, Y., MEHMOOD, R., KIMOTO, C., FUJITA, T., TOUMA, K., KANEKO, A., KAMIKAWA, Y., MORIYAMA, T., YANAGIDA, T., KANEKO, H. and YONEDA, Y. (2013). Importin alpha subtypes determine differential transcription factor localization in embryonic stem cells maintenance. *Dev Cell,* 26**,** 123-35.

YEGANEH, M., SEYEDJAFARI, E., KAMRANI, F. A. and GHAEMI, N. (2013). RNA-binding protein Rbm47 binds to Nanog in mouse embryonic stem cells. *Mol Biol Rep,* 40**,** 4391-6.

YING, Q. L., NICHOLS, J., CHAMBERS, I. and SMITH, A. (2003a). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell,* 115**,** 281-92.

YING, Q. L., STAVRIDIS, M., GRIFFITHS, D., LI, M. and SMITH, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol,* 21**,** 183-6.

YOUNG, R. A. (2011). Control of the embryonic stem cell state. *Cell,* 144**,** 940-54.

YU, H. B., JOHNSON, R., KUNARSO, G. and STANTON, L. W. (2011). Coassembly of REST and its cofactors at sites of gene repression in embryonic stem cells. *Genome Res,* 21**,** 1284-93.

YUAN, H., CORBI, N., BASILICO, C. and DAILEY, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev,* 9**,** 2635-45.

YUE, Z., MALDONADO, E., PILLUTLA, R., CHO, H., REINBERG, D. and SHATKIN, A. J. (1997). Mammalian capping enzyme complements mutant Saccharomyces cerevisiae lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc Natl Acad Sci U S A,* 94**,** 12898-903.

ZAN-KOWALCZEWSKA, M., BRETNER, M., SIERAKOWSKA, H., SZCZESNA, E., FILIPOWICZ, W. and SHATKIN, A. J. (1977). Removal of 5'-terminal m7G from eukaryotic mRNAs by potato nucleotide pyrophosphatase and its effect on translation. *Nucleic Acids Res,* 4**,** 3065-81.

ZENKLUSEN, D., VINCIGUERRA, P., WYSS, J. C. and STUTZ, F. (2002). Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol Cell Biol,* 22**,** 8241-53.

ZHANG, X., HUANG, C. T., CHEN, J., PANKRATZ, M. T., XI, J., LI, J., YANG, Y., LAVAUTE, T. M., LI, X. J., AYALA, M., BONDARENKO, G. I., DU, Z. W., JIN, Y., GOLOS, T. G. and ZHANG, S. C. (2010). Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell,* 7**,** 90-100.

ZHOU, J. M., XING, F. Y., SHI, J. J., FANG, Z. F., CHEN, X. J. and CHEN, F. (2008). Quality of embryonic bodies and seeding density effects on neural differentiation of mouse embryonic stem cells. *Cell Biol Int,* 32**,** 1169-75.